January 2020

Impacts Of Market Forces And Agricultural Practices On Land Surfaces Related To Scientific Modeling And Remote Sensing Applications

Jon Starr

IMPACTS OF MARKET FORCES AND AGRICULTURAL PRACTICES ON LAND SURFACES RELATED TO SCIENTIFIC MODELING AND REMOTE SENSING APPLICATIONS

by

Jon Paul Starr
Bachelor of Science, University of Wyoming, 2010

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota
August
2020

Name: Jon Starr

Degree: Master of Science

This document, submitted in partial fulfillment of the requirements for the degree from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

DocuSigned by:
Jianglong Zhang
2787AED1C19A41B...
Jianglong Zhang

DocuSigned by:
Haochi Zheng
17774751C7844B3...
Haochi Zheng

DocuSigned by:
David C. Roberts
00ADC3E54A8A450...
David C. Roberts

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

DocuSigned by:
Chris Nelson
2E0AF088C733403...
Chris Nelson
Dean of the School of Graduate Studies

7/27/2020
Date

PERMISSION

Title IMPACTS OF MARKET FORCES AND AGRICULTURAL PRACTICES ON LAND SURFACES RELATED TO SCIENTIFIC MODELING AND REMOTE SENSING APPLICATIONS

Department Atmospheric Sciences

Degree Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Jon Starr
07/24/2020

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I wish to express my sincere appreciation to the members of my advisory Committee, Dr. Jianglong Zhang, Dr. Haochi Zheng, and Dr. David C. Roberts, for their guidance and support during my time in the master's program at the University of North Dakota.

To my mother, Annette, and my grandfather, Earl, who gave the encouragement and guidance needed to complete this task.

ABSTRACT

Due to the rotational needs of crops, homogenous crop fields, and external influences such as market and policy changes, crop production generates significant changes to the landscape on annual and semi-annual basis. In this study we looked at two aspects of this change.

In the first aspect of the study, we attempt to account for market and policy driven producer's decision making through a new model constructed by pairing an economics model with the ALMANAC crop simulation model via a two-way coupling. This coupled model approach integrated farmer's land-use choices based on relative economic returns and produced dynamic land use probabilities for ALMANAC simulations through a feedback loop. The coupled model approach was inter-compared with static crop modeling through a historic acreage approach, and comparable accuracies were found from both modeling efforts for the 2014 growing season. Furthermore, as a proof-concept effort, the method was applied to evaluate the impact of two scenarios on crop simulations: major crops (maize, soybean, and wheat) intensification through price increases (e.g. market change), as well as incentivized grassland conservation (e.g. policy change). The results of this sensitivity study suggest that the coupled system has the capability of integrating economic factors into traditional crop simulation, allowing for insight into the impacts of changes in markets and policies on agricultural landscapes and crop yields.

In the second aspect of this study, changes to surface albedo driven by these landscape changes are investigated. Using collocated Moderate Resolution Imaging Spectroradiometer (MODIS) derived Bidirectional Reflectance Distribution Function (BRDF) with the Cropland Data Layer (CDL), we computed the daily albedo of homogenous agricultural fields across the United States for 55 crop types by wavelength, sky-type, day of year, crop, and hardiness zone over a four-year period (2015-2018). This study suggests that cropland spectral albedo is complicated by large variations over the course of the growing season, which can result in changes in reflectivity up to a factor of 2 at most wavelengths. This change was found to be unique per crop type, but predictable year-to-year for individual crops within specific regions, so generating a lookup table that incorporates these factors for use in remote sensing and atmospheric modeling applications is viable for albedo estimation. Additionally, impacts of crop types on broadband albedo were studied and found to be less conspicuous than the individual wavelength counterpart, but still significant over cropland. The results were used to evaluate the accuracy of a common method of albedo estimation, where NDVI is used as a proxy for albedo over cropland, and the NDVI method was found to have some significant limitations dependent on wavelength and day of year. Finally, a database of surface albedo variations as a function of growing period is constructed for 55 crops common to croplands across the United States. The constructed database can be used to aid both satellite remote sensing applications and long-term weather modeling efforts by providing a method for parameter adjustments based on crop driven albedo changes, including changes in cropland composition related to commodity markets and other external factors.

CHAPTER 1

INTRODUCTION

Due to the high acreage demands required for food production, much of the land-surface of earth, that is productive, is utilized as cropland, leading to large areas on homogeneous land surfaces. However, this cropland is undergoing changes in appearance, soil coverage, and ecosystem functionality at both seasonal and annual timescales due to the cyclical nature of crop growth. This can lead to complications in both modeling and remote sensing algorithms which require these properties as part of their equations.

For example, to promote soil health and increase yields crop producers change crop selection in a rotation on, typically, an annual basis. This rotation can follow a pattern defined by standard practices or can be improvised by the producer's evaluations and reactions to market changes, weather trends, risk management, as well as other factors. As a result, the crop present at any given location varies at annual timescales in a challenging to predict manner.

To study the potential impacts on yields and soil health of these crop selections, along with associated agricultural practices and management techniques, crop simulation models are frequently employed. These simulation models, such as ALMANAC or EPIC, typically work on a daily timestep, calculating the crop growth, stages, and demands throughout the growing season. Meteorological inputs, such as precipitation, temperature, wind, and humidity, as well soil inputs, such as water holding capacity and porosity, are combined in the model to calculate biomass gain, crop maturity, stress factors, yields, and soil impacts. These crop simulations can further be scaled to decades or centuries to investigate long-term impacts and incorporate multiple scenarios, climates, or management practices, allowing for the investigation into a wide variety of agricultural research

questions. However, while these models can simulate crop rotations, and are often utilized in fixed, standard rotations, the impacts of the producers decision making is not typically incorporated, particularly with the sophistication to account for feedback effects of large-scale market changes.

Additionally, the changes/rotations of crop types can have a significant impact on the surface albedo at a given location, due to the disparity in both maturity timing and overall reflectivity of the planted crops. Impacts to this area are essential to ascertain as albedo is a critical input variable into several atmospheric science research areas, including atmospheric models and remote-sensing based instruments. Studies that incorporate these tools often deal with these vegetation induced albedo changes in a static manner, or through generic assumptions, and as a result albedo changes cause by specific crop coverage are not accounted for.

In this thesis we strive to tackle some of the questions brought from the constantly changing nature of cropland through two different approaches: first, we attempt to model the interaction of market forces with the crop landscape, and the subsequent impacts to yields through a coupled, two-way interaction between a crop simulation model and an economics model; secondly, we investigate the impacts of the changing agricultural landscape to surface albedo through satellite-based observations over a multi-year period. Through these methods we hope to answer two main questions: 1) can an economics and crop-simulation coupled model be constructed, and if so what impacts to the landscape will be observed?; and 2) does the albedo of cropland across the United States change significantly across crops and regions, and if so, can the results from this study be utilized to construct a crop-focused albedo database?

CHAPTER 2

DATASETS

In this chapter, datasets and models used are discussed in details for 1) modeling and studying of the two-way interactions between crop models and an economics model (the first part of this study as illustrated in chapter 3); 2) the study of changes in narrowband and broad band surface albedo during the plant growth period (the second part of the study as shown in chapter 4).

2.1 The SSURGO database

The SSURGO database is a spatially referenced database containing soil profile and general characteristics information for the majority of the United States land area at a scale varying from 1:12,000 to a max of 1:63,360**.** This dataset was chosen as the soil input for crop modeling in the first section of the study due to its extensive scope and high spatial resolution as the proper fit to simulate the prairie pothole area at a fine spatial scale. Internally the ALMANAC model utilizes the SSURGO information to generate soil profiles based on water holding capacity, soil depth, and chemical components (Kiniry, et al. 1992). The database contains a collection of uniquely identified soil types covering the whole of North Dakota, with each soil characterized by the depth of each layer of the profile as well as the overall properties of the soil, expressed in means and ranges, for each independent layer of soil. Typically, these soil files are divided into separate databases for each county or distinct geographic region. For the Prairie Pothole Region included in the

first part of the study (chapter 3) there are a total of 6,995 unique soil profiles (median area of 328 ha) broken into 39 unique databases with a scale of 1:12,000. The SSURGO dataset was acquired from the NRCS data gateway website (https://datagateway.nrcs.usda.gov/) (Soil Survey Staff 2016).

The non-irrigated Land Capability Class (Nirr LCC) in SSURGO database, which defines a soil's potential for crop production during standard rain-fed farming practices, is used to determine the soil productivity in the first part of the study. Soils in North Dakota fall into the range of LCC 2 to LCC 8 classifications; these groups have increasing levels of limitations on crop growth, with LCC 2 containing the least, LCC 8 the most, which reduce overall potential productivity to varying degrees. (Soil Conservation Service, USDA 1961). LCC is used in the first part of the study to identify the potential impacts of migration patterns in land uses as demands for crops increase or decrease total acreages over finite resources.

2.2 The North American Regional Reanalysis (NARR) data

The North American Regional Reanalysis database, used in the first part of the study, is generated by combining research weather models with past observations to complete a gridded summary of the local atmospheric conditions at resolutions up to 32 km per grid. This data is provided in eight-times daily and daily summary formats at defined model pressure levels. For the Prairie Pothole study region (for the first study as described in chapter 3) the grid spacing is on average 32.40 km, +/- 0.05km. While this resolution is coarser than other available datasets, such as the 4km PRISM dataset (Daly, Taylor and Gibson 1997), this system was chosen due to the similar lineage to data

generated from current generation climate models. This allows for past, present, and future climate simulations to be run without recalibration of the crop models when utilizing the same assumptions. The NARR dataset provides a wide variety of meteorological variables, such as wind speed and temperature. For the first part of the study temperature at the 3-hour time step scale was used, while precipitation, wind speed, relative humidity, and solar energy inputs were derived from the daily summaries. The ALMANAC model uses the solar radiation, temperature, and precipitation to calculate growth rates and stresses, while wind speed, relative humidity, and solar radiation are used to determine potential evaporation. The NARR dataset was acquired from the Earth Systems Research Laboratory (https://www.esrl.noaa.gov/psd/data/gridded/data.narr.html) (Mesinger, et al. 2006).

2.3 The Cropland Data Layer data

In this study the Cropland Data Layer (CDL) is used to provide annual locations and crop identifications for all cropland and other land uses within the United States at a 30m resolution. For the first part of this study (chapter 3) we utilized the CDL to map the seven major crop types: spring wheat, maize, soybean, oats, sunflower, canola, and alfalfa inside the Prairie Pothole Region, but accounted for other crops and land uses given by the CDL in our final land use area estimates. Within the Prairie Pothole Region, a total of $1.5 \cdot 10^5$ ha or 74.7% is in land cover accounted for in the first study, leaving 25.2% of the area consisting of wetlands, other non-farmable, and non-study crops which are held static. For the second part of the study (chapter 4), to calculate the impact of crop type on surface albedo, precise information about the crop composition present within each MODIS pixel is required. The CDL is an annually produced georeferenced raster file that defines surface

crop types for the majority of the United States at a resolution of 30 to 56 m. Crop types are determined through analysis of satellite imagery (Boryan, et al. 2011). The data is then processed and verified against the National Agricultural Statistics Service (NASS) and the Farm Service Agency (FSA) farmer records to increase accuracy of the detection algorithm (Boryan, et al. 2011). A total of 132 land use types are included in the CDL data, of which a total of 47 are found within the study region. In this study The CDL dataset was obtained from the Cropscape web site (https://nassgeodata.gmu.edu/CropScape/) (USDA-NASS 2016).

2.4 The ALMANAC model

In the first part of the study, the USDA's ALMANAC model is selected for use in crop simulation for its inherent connections with the Soil Survey Geographic Database (SSURGO) (Soil Survey Staff 2016) database, its ability to accurately simulate a wide range of crops, the depth of its field management options (Xie, Kiniry and Nedbalek, et al. 2001), as well as the extensive reviews on the input sensitivities that have been completed, such as Xie, Kiniry and Williams (2003). The ALMANAC model is a daily time step crop simulation model originally based on the EPIC model (Mearns, Mavromatis and Tsvetsinska 1999). The ALMANAC model produces a point based, soil specific simulation of the growth, health, and yield of a variety of crops including the seven selected crops as mentioned in the introduction section for this study. Additionally, the ALMANAC model was chosen due to its ability to simulate at a per-soil basis, directly matching the input of the economics model, allowing for the investigation of land use migration at this same level.

The ALMANAC model requires three main inputs: management protocols, soil characteristics, and local meteorological conditions. For management, planting and harvesting dates from statewide climatological averages are used for each study crop, fertilizer applications are static and set to once at planting if required by the crop, no irrigation or other in-season intervention is included. Soil characteristics and components are handled internally through the ALMANAC model utilizing the SSURGO soil dataset dated 2015. Meteorological information derived from the NCEP North American Regional Reanalysis (NARR) dataset. Each SSURGO soil area in the study region is geometrically subsected by the native NARR grid spacing of 32km using geometric intersection and treated independently, resulting in 18,136 individual simulations for each crop with a median area of 119.7 ha. For this study, each simulation was run for the specified year after a one-year spin-up; with longer time frame spin-ups tested for this study but no major changes in results were found.

2.5 The economics land use model

The individual-based economics land-use model focuses on the agricultural profitability of producing different crops under policy and market assumptions (Kharel, Zheng and Kirilenko 2016). The spatially-explicit land-use model calculates net return of each crop and determines the crop composition for a given unit using crop yields simulated by the ALMANAC model. The net return of a soil unit s ($s = 1 \dots S$) in year t, to be assigned for a certain use or grow a particular crop c ($c = 1 \dots C$), is calculated as $\pi_{s,c,t} = P_{c,t} Y_{s,c,t-1} - C_{c,t}$, where $Y_{s,c,t-1}$ is the crop yield simulated by the ALMANAC model for a particular soil type and productivity in year

$t-1$, $C_{c,t}$ is the crop production cost, and $P_{c,t}$ is the expected price for a crop c in year t. We assume that individual landowner estimates the expected economic return to grow a certain crop based on the previous observations, knowledge of the soil type, productivity of the land, as well as the current price information of future market movements. Therefore, the likelihood of growing certain crop in a given unit is determined by the relative profitability of that crop compared to other competing land-use alternatives, by assuming that each landowner makes optimal choices to maximize the total economic return.

In this modeling exercise, we simplified the management details and used static managements in ALMANAC by assuming that farmers grow certain crop under a general fixed management scheme. However, in reality, the choices of management practices as well as their costs likely affect farmer's decision making on land use and crop type selections. The profit maximization, therefore, can be further achieved by modeling farmer's management choices based on physical conditions related to soil and climate as well as economic factors such as the fluctuations of input prices.

Table 1: Direct Production Costs and Economic Net Returns of Crops

	Corn	Soybeans	Wheat	Oats	Sunflowers	Canola	Alfalfa	Grass	Forest
Direct cost ($/acre)[1]									
Seed	87.63	68.64	19.09	12.75	39.06	54.34	2.93		
Fertilizer	111.95	15.35	67.96	37.81	57.73	72.00	5.80		
Crop chemicals	22.79	22.48	34.73	14.97	43.47	30.30	0.93		
Crop insurance	22.08	17.84	15.61	14.55	12.88	12.97	5.17		
Fuel & oil	27.26	19.21	17.62	15.46	18.87	22.11	12.50		
Repairs	31.07	19.98	18.76	19.22	20.45	21.67	17.20		
Custom hire	5.05	4.24	7.58	14.28	11.88	5.32	2.15		
Land rent	49.34	53.91	34.62	17.59	34.22	28.99	20.31		
Operating interest	6.89	4.34	4.13	3.29	5.02	4.23	3.11		
Miscellaneous	1.78	0.95	1.58		0.93	0.29	0.60		
Drying expense	9.84								
Storage	0.76								
Crop prices ($/unit)[2]									
Baseline	3.70	10.10	5.99	2.58	20.10	16.73	74.04		
"Crop Intensive"	4.44	12.12	7.19	2.58	20.10	16.73	74.04		
Net returns ($/acre)[3]									
Baseline	45.32	55.97	28.79	2.33	20.10	16.73	12.59	4.76	3.03
"Crop Intensive"	56.65	66.04	34.61	2.33	20.10	16.73	12.59	4.76	3.03
"Grassland Incentive"	45.32	55.97	28.79	2.33	20.10	16.73	12.59	44.76	3.03

[1] The direct costs of crop productions were collected from FINBIN, the FINPACK financial database (https://finbin.umn.edu/Home/AboutFinbin) for the ND state in 2014.

[2] The crop prices are in $/bushel for corn, soybeans, wheat, and oats, in $/cwt for sunflowers and canola, and in $/ton for alfalfa/hay. Same crop prices were used in the Baseline and "Grassland Incentive" scenario.

[3] The economic net returns of grassland and forestland were adopted from Lubowski et al. 2006, 2008.

For the first part of the study, we treated prices as exogenously determined outside of the system based on the fact that the study region is relatively small and has played a moderate role as a "price-taker" in the domestic commodity markets. We collected crop price information for small crops (oats, sunflowers, canola, and alfalfa/hay) and production cost data, shown in table 1, for all seven crops in the study region from the Farm Financial Data Base (http://www.finbin.umn.edu) hosted by the Center for Farm Financial Management at the University of Minnesota. We chose this dataset to utilize the real-world farmers' budgetary information by considering farming itself is a systematic decision-making process and each management choice is made in conjunction with the others. We imported

annual 2014 market year prices from USDA National Agricultural Statistics Services (NASS) (USDA-NASS 2018) for the major crops to reflect the general fluctuations across domestic markets. For scenario analyses, we used a different set of prices to demonstrate the potential increases in market demand for the major crops. We used FINBIN prices information for the small crops given these prices are likely determined by the regional/local market. To focus on modeling land productivity for crop production, we simplified the modeling of grassland and forestland by using a static average net return reported by Lubowski (2006, 2008) based on the spatial association of individual soil unit with each ND county and adjusted for inflation. The final estimated net return of each land use alternative was transformed to a probabilistic surface using logistic distribution to represent the likelihood of the land-use transition (Lubowski, Plantinga and Stavins 2006; Lubowski, Plantinga and Stavins 2008; Lewis and Plantinga 2007).

2.6 MODIS BRDF Albedo Product:

The MODIS Bidirectional Reflectance Distribution Function (BRDF) / Albedo Parameter level 3, 500m gridded, dataset (collection 6) is used in the second part of the study (chapter 4) for estimating surface albedo at seven wavelengths (0.47, 0.56, 0.65, 0.86, 1.24, 1.64, and 2.13μm). This dataset calculates the effective black- and white-sky albedo parameters using a kernel driven approach over a 16-day moving window (Schaaf, et al., 2002). Through black- and white-sky albedo the surface albedo can be calculated for specific solar zenith angle and atmospheric scattering ratio, from (Lucht, et al., 2000)

as follows:

$$\alpha = \left(1 - SR_{dif}\right) \cdot \alpha_b + SR_{dif} \cdot \alpha_w \quad (1)$$

where α is the total albedo, SR_{dif} is the ratio of indirect to direct sunlight reaching the surface, α_w is the white-sky albedo, and α_b is the black-sky albedo. This dataset further includes quality control flags for both the overall pixel quality and the snow coverage of any given pixel within the area during the 16-day moving window. Throughout this study, only snow/ice-free "good" or "best" quality pixels were utilized to focus on changes seen to the land surface due to crop growth and/or producer management.

The MODIS BRDF dataset additionally provides a calculated broadband albedo for three spectral ranges, Visible (300-700nm), IR(700-5000nm), and Shortwave(300-5000nm), derived from the reflectance of the individual channels as shown in equation 2 (Lucht, et al., 2000):

$$\alpha_r = \sum_i w_i \alpha_i \quad (2)$$

Where the α_r is the albedo of the spectral range, α_i is the albedo as calculated for each independent channel, and w_i is the weighting factor for each channel. w_i was calculated through computational means to derive the relative impact of each channel to the overall albedo.

2.7 Plant Hardiness Zone

To account for changes in growing seasons the impacts of each zone to albedo in the second study the selected locations are separated into specific hardiness zones (HZ)

using the USDA's Plant Hardiness Zone (PHZM) defined zones (USDA-ARS, 2012). The PHZM zones are split into half-zone divisions of 2.8°C defined through the calculation of the 30-year average extreme minimum temperatures of 7,983 North American weather stations. These are then interpolated, with elevation impacts incorporated, into the final map (Daly, et al., 2012). However, in order to streamline the categories in the final database, this study combines the PHZM half-zone divisions into their full zone (5.6°C) counterparts.

By accounting for the growing season length and temperature, the calculations of crop albedo can be grouped by commonality of growth cycles. This is possible as the local climate largely defines both the planting and harvest dates, along with growth rate of particular crops, resulting in common growth patterns within a given HZ. Grouping by HZ could potentially lead to reduction in uncertainty due to plants of the same type being planted and harvested near the same time within each HZ, thus having a similar growth profiles throughout the year.

2.8 MODIS Reflectance data

To aid in the application of using NDVI as a proxy for per-channel albedo, level 2 MODIS surface reflectance (MOD09 + MYD09) (Vermote, 2015) was utilized. Data was extracted from both Aqua and Terra satellites at a 500m resolution for the 2013 season, giving two independent reference points each day for calculating NDVI. The NDVI calculation utilizes the red (645nm center) and near infrared (860nm center) channels to determine the vegetation index as follows (Rouse Jr, et al., 1973):

$$NDVI = \frac{(REF_{NIR} - REF_{Red})}{(REF_{NIR} + REF_{Red})} \quad (3)$$

where REF_{NIR} is the reflectance of the near infrared MODIS channel (860nm) and REF_{Red} is the reflectance of the red MODIS channel (650nm). The resulting unitless index ranges from -1 to 1. NDVI allows for the analysis of plant growth and activity due to the increased absorption of the red channel and increased reflectance, relative to the underlying soil, of the near infrared channel of the chlorophyll active within plants. NDVI values > 0.4 are generally associated with levels of plant growth with especially high values (>0.8) possible during times of peak growth of field crops. Finally, for the investigation into NDVI as a proxy of albedo, both the NDVI and albedo pixels were masked using the MODIS annual land cover in place of the CDL layer, allowing for both single and multi-crop MODIS pixels to be included in this aspect of the study.

CHAPTER 3

Evaluating sensitivities of economic factors through coupled Economics-ALMANAC model system

3.1 Rationale

To accurately estimate crop yields at regional scales for economic analysis and prediction crop models have been developed and used in simulating crop yields and soil health over selected regions (Williams, Renard and Dyke 1983; Hertel and Rosch 2010). Some crop models commonly used for this simulation include Environmental Policy Integrated Climate (EPIC) model (Williams, Renard and Dyke 1983); the Agricultural Land Management Alternative with Numerical Assessment Criteria (ALMANAC) model (Kiniry, et al. 1992); the Decision Support System for Agrotechnology Transfer **(**DSSAT) (Jones, et al. 2003); the Agricultural Production Systems sIMulator (APSIM) (Keating, et al. 2003); and Crop Environment REsource Synthesis (CERES) (Ritchie and Otter 1985).

However, while these crop simulation models take into account both weather and soil changes, one factor lacking in the above crop models is the dynamic impact of land use changes due to economic factors, such as market fluctuations and changes in policy. These factors can influence landowner's decision-making on land uses and management practices, and thus further affect crop yields. The inherent agricultural productivity of land is determined by its biophysical characteristics and the surrounding climate, making inputs in these areas near static at annual time scales. However, decisions on land-use and management practices are dynamically driven by the individual landowner and can

change at an annual time scale based on the economic return from each available alternative. Studies at the regional scales often simulate crop yields with a set of fixed assumptions on land uses and management practices throughout the analyses. However, this process ignores the impact of landowner's dynamic decision-making at a local scale with multiple soil profiles as responses to changes in local economic conditions, such as changes in market and policy conditions, reflected by crop prices and policy incentives. Furthermore, over decadal time scales, the policy-induced changes in land allocation and in farming practices for crop production will, in turn, impact soil health and its agricultural productivity at longer time scales, which feeds back into the decision-making process.

While coupling crop models with economic models to adjust for this impact has been the topic of several studies, these studies are typically setup in a one directional fashion with results from the economic model feeding the crop simulation model (e.g. Briner, et al. 2012; Robertson, et al. 2012) or the crop simulation model results feeding the economic model (e.g. García-Vila and Fereres 2012) without a two-way interaction as attempted in this study. Additionally, while looking at long term changes, crop production simulations of responses to future scenarios often either utilize gridded data, include few locations, or utilize a single soil profiles per location (White, et al. 2011). Therefore, without taking into account the feedback loop between soil health and economic decision-making at a finite, individual soil-based resolution, the simulation results from the traditional crop modeling approaches typically ignore the two-way interaction between annual yields and profits and the resulting land use changes. This shortcoming likely leads

to a tendency to move towards unrealistic depictions during these longer-term simulations.

In this study, we expand upon the standard crop simulation model paradigm by generating dynamic agricultural land use choices and implementing them into large scale crop simulations. We accomplish this through a two-way linked economics land use model and crop model at an annual time step looking at seven crops common to the study area, maize (*Zea mays*), soybean (*Glycine max*), spring wheat (*Triticum aestivum*), oats (*Avena sativa*), alfalfa (*Medicago sativa*), canola (*Brassica napus*), and sunflower (*Helianthus annuus*). In this proof of concept study, we focus our preliminary studies on the following questions:

(1) Can a linked system be developed to incorporate economic factors in modeling land-use at soil-based resolution and offer advantages in crop simulations?

(2) Can we use the linked economics-crop modeling system to further evaluate the sensitivity of economic factors, such as policy and market changes, on crop yields and soil health prediction?

3.2 Data and Models

The selected area for the study is the Prairie Pothole region of North Dakota (Figure 1), a region spanning east and north of the Missouri River with extensive grassland and wetland coverages for providing crucial habitats for endangered species and other ecosystem services. This region was selected for its well-known high soil productivity as well as the recent significant grassland conversion to corn and soybean cultivation (Ojima, et al. 2002, Wright and Wimberly 2013).

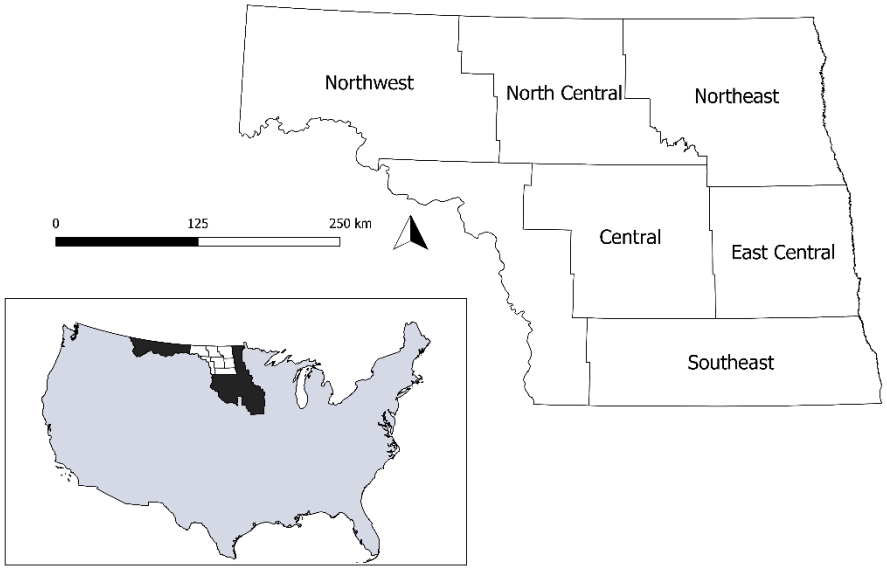

Figure 1: Locations included in this study and their relative positioning inside the United States Prairie Pothole Region.

For the simulations three primary environmental datasets were used. The structure and properties of the soils in the study region were obtained by using the 2015 version of the Soil Survey Geographic database (SSURGO) (Soil Survey Staff 2016) covering the state of North Dakota. Meteorological variables were acquired through the North American Regional Reanalysis (NARR) (Mesinger, et al. 2006) dataset from the National Center for Environmental Prediction. Finally, the Cropland Data Layer (CDL) (USDA-NASS 2016) was used to determine the historic crop locations and total area. In addition to these datasets, an economics framework (Kharel, Zheng and Kirilenko 2016) and the ALMANAC crop model (Kiniry, et al. 1992) were linked together, referred to as ALM-EC through the rest of the study, and are applied in this study.

3.2.1 Model Calibration

Before the study the ALMANAC model was calibrated for the study region by adjusting the built-in crop parameters within the model to match local crop varieties. This calibration step is needed as varieties of commonly grown crops differ from region to region due to the specific needs or limitations of each area. To compensate for this impact on the overall growth and eventual yield each crop requires separate calibration to the ALMANAC parameters.

Calibration was completed for the study crops utilizing annual yield as the primary factor for a single year. A randomized set of points (set so as $n > 1000$ per crop) was generated within the state of North Dakota based on the estimated crop grown within that year at that location as reported by the CDL (Figure 2). These were then repeated for each year through the 2001-2013 growing seasons using a similar technique, each year individually simulated, with spin-up period, but with a shared point selection filtered to only include those locations with frequent reoccurrence of the selected crop. The county by county yield aggregates were tabulated and compared to the NASS given statistics for that county in that year. As an example, Figure 2 shows the validation points selected to compare model-simulated yields with the reference county reported yields for spring wheat, maize, and soybean respectively. The primary parameters adjusted included total growing degree days to account for the shorter growing season as well as increased water stress tolerance to compensate for both the climate of the area as well as the diffuse nature of the precipitation in the weather model data used in this study; additional minor growth parameters for each crop were adjusted as needed. The parameters were calibrated until the

resulting simulation annual county level means were within 10% of NASS reported annual mean yields at the county level.

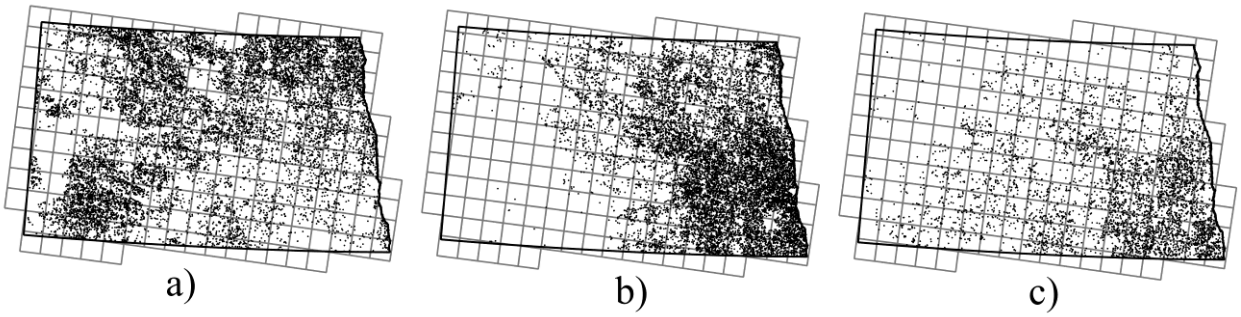

Figure 2: Locations used to simulate spring wheat (a), soybean (b), and maize (c) during model calibration over North Dakota with the native NARR weather data grid superimposed.

3.3 Methods and Experimental design

In this study, we developed a linked crop-economics model and inter-compared the performance of the linked crop-economics model with crop simulations from a static crop modeling approach. After completion of this initial stage, the linked crop-economics model is used, as a concept proofing effort, to investigate the sensitivity of crop simulations with respect to major crops intensification as well as grassland conservation scenarios as a proxy for perturbations in market and policy conditions respectively.

3.3.1 Description of the economics and crop model feedback loop

A primary goal of this study is developing a linked economics land-use and crop model system which integrates economic-based land use changes into the crop simulation modeling process. To facilitate this, we established the looped-feedback pattern as described in Figure 3, interconnecting the models directly. At the beginning of the process, weather, soil, and crop datasets for the previous year are used as inputs to the ALMANAC model for estimating yield performance for all crop and soil combinations within the study area for the previous year. The yields are used as inputs, along with policy and markets information, to the economics model. The economics component calculates the economic

return of each land-use alternative and the relative likelihood of each crop being grown in a specific location with a unique combination of soil and weather for the current year. In other words, the land-use probabilities from the economics land-use model can further prescribe the allocation of land within a particular unit. This structure demonstrates farmer's decision-making process uses simulated crop yields in previous year as a reflection of farmer's knowledge or observation of land productivity as well as the current prices as a proxy of market and policy changes. With the detailed information of crop allocations within in each individual soil unit, ALMANAC model produces the final yields over those areas and soil information for the study region for the current year. Each subsequent (annual) time step repeats the entire process save for the calibration stage, using the soil state from the previous year's simulation. As an example of this connection, Figure 3 shows this study's concept proof design, focused on simulating land use probabilities for the year 2014 (as shown in Figure 4). Note that in this practice, yields from 2013 are used for predicting agricultural land use and land change for 2014, using price and policy as control variables. This provides a potential forecasting capability for future studies.

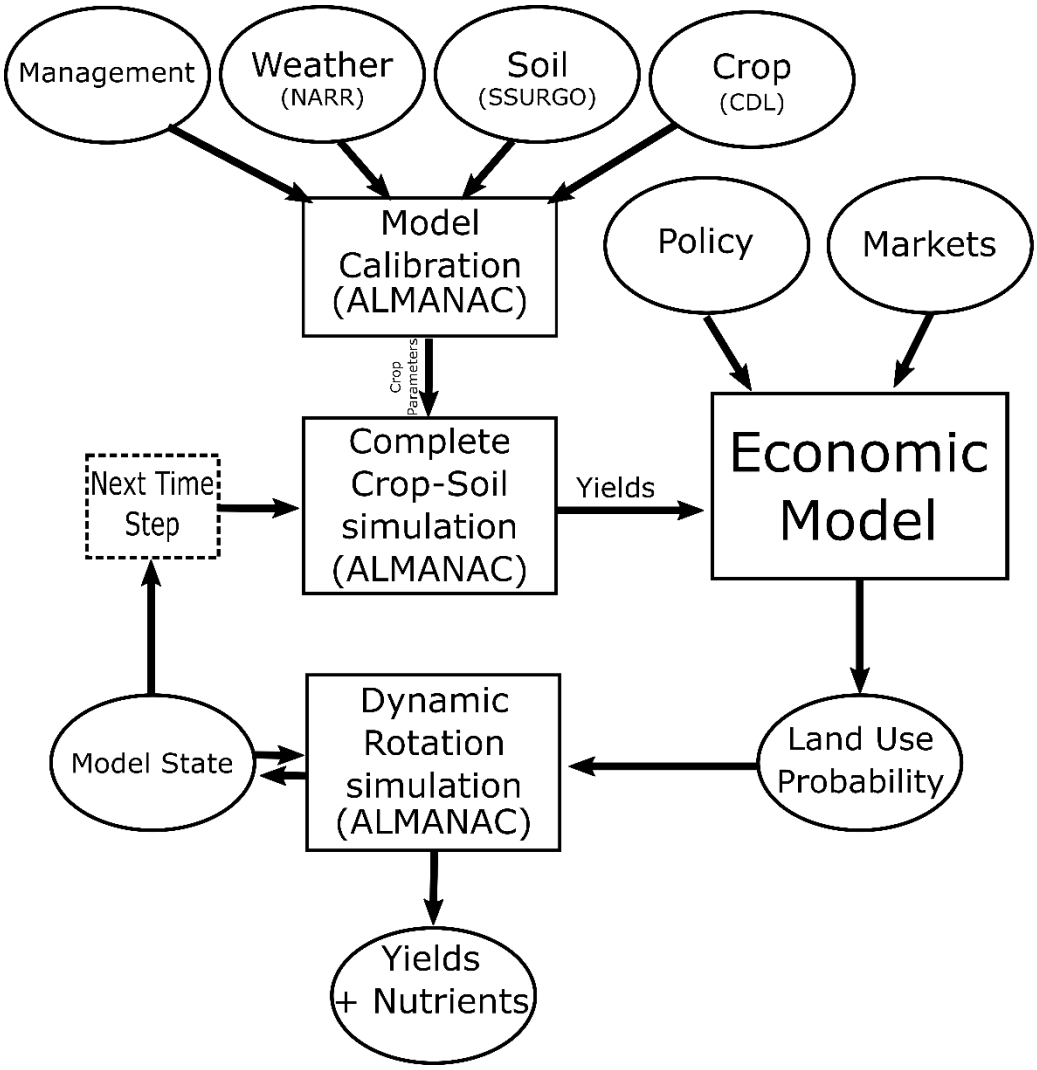

Figure 3: Study workflow for crop simulation using land use probability from the economics model.

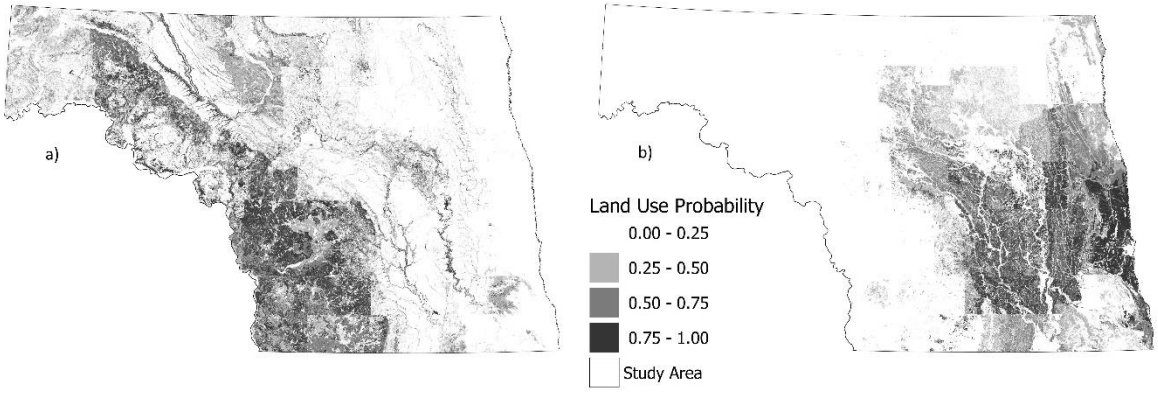

Figure 4: Resulting land use probabilities for a) grassland and b) soybean generated by the economic model for the 2014 season.

3.3.2 Evaluation of land use area and yield simulated using the linked economics-crop model system

Crop simulations without (section 3.3.3; standalone crop simulation) and with (section 3.3.4; linked crop-economics model) the use of the economics land-use model are performed and inter-compared for year 2014, setting the basis for evaluating the influence of crop and market changes on crop simulations as laid out in section 3.4.

The standalone crop simulation approach follows a simplified process that determines the composition of land uses for 2014 based solely on historical records of crop patterns observed over the period of 2010-2013, at the coarsest resolution (30 m) the CDL supplies during that timeframe. The linked crop-economic model approach utilizes a paired crop simulation and economics land-use model to determine the most likely cropping patterns in the 2014 season, emphasizing the two-way linkage as a more systematic method. The results from each of these two approaches were used in the simulations of the 2014 growing season for comparison. Crop yields and land use simulations from both approaches were then compared with both the estimated land uses as described by the 2014 CDL as well as actual yields reported by NASS.

3.3.3 Crop simulation with historical crop patterns

As seen in the flowchart of Figure 5, the standalone model approach used the ALMANAC model to simulate crop yields for the 2014 season using the fixed historical crop percentages. To determine the historic crop planting percentage utilized in each soil type, the CDL from 2010-2013 was geospatially intersected with the individual SSURGO soil types to determine the most commonly seen crops for each soil area. These were then

applied to determine the soil and crop combinations for the ALMANAC model and run for the 2014 growing season. The resulting yields were tabulated at an individual SSURGO soil type calculating areas using the CDL's historic area percentage of each crop in that soil type for direct comparison to the ones generated by the economics land-use model.

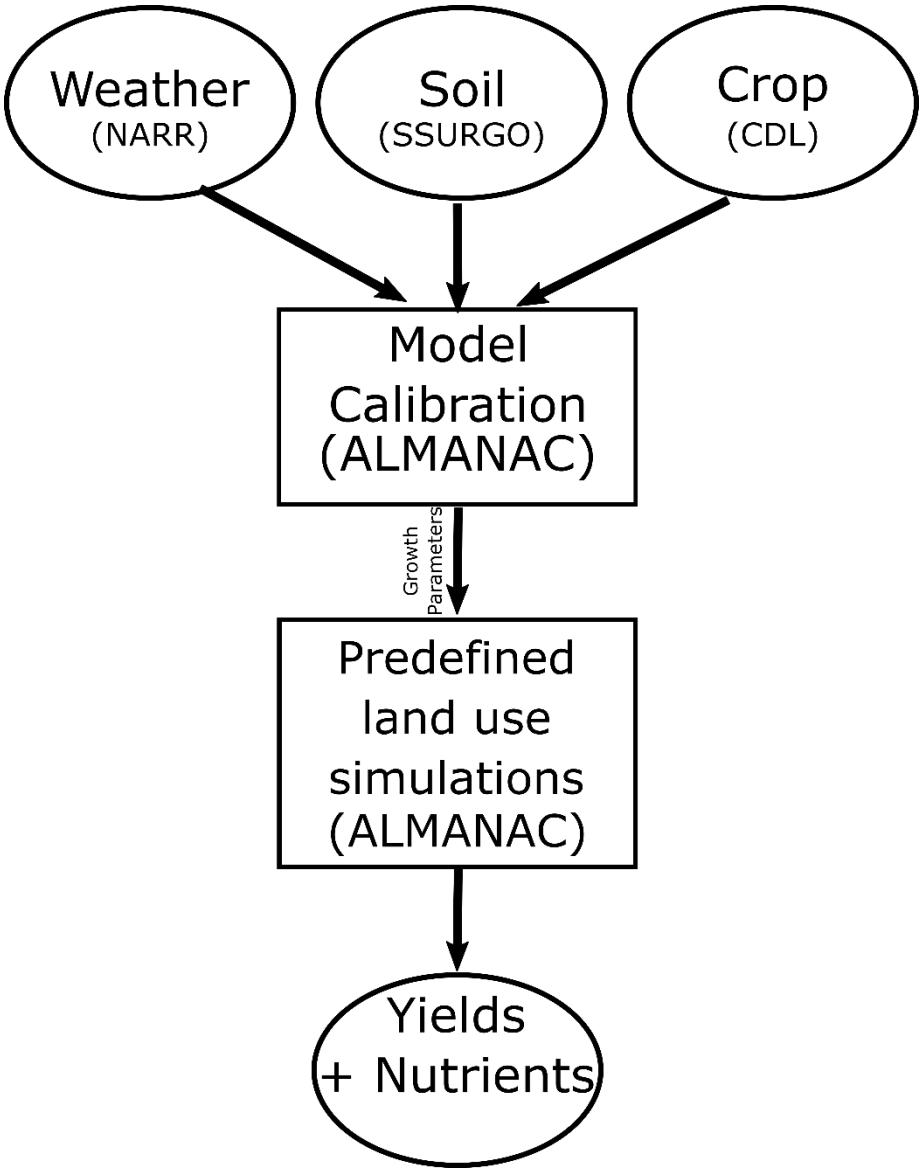

Figure 5: Study workflow for the non-coupled, standalone crop simulations using historical acreages

3.3.4 Crop simulation with land use probability from the economics model

As a comparison, the second approach simulates crop yields using land use probability prescribed by the ALM-EC model. CDL, SSURGO, and NARR data from 2013 are used as inputs for the ALMANAC model which simulates soybean, maize, spring wheat, sunflower, canola, oats, and alfalfa yields for all soil map units over North Dakota for 2013 (Figure 3). These simulated crop yields for 2013, along with crop prices and management costs determined by a specific scenario, are used as inputs for the economics model. The land use probability (Figure 4), as the outcome from the economics model, is utilized to generate possible soil-crop combinations for ALMANAC to simulate for the 2014 season, with the resulting total land area and production for each soil-crop simulation weighted by the economic probability per soil.

To test the integrated approach via coupling the two models, the simulated land use composition was then compared with the estimated acreage derived from the previous years' CDL as well as the CDL estimated 2014 crop acreages.

3.3.5 Evaluation of the impacts of market price and policy changes on crop simulations

Upon evaluation of the linked crop and economics modeling system, we extended crop simulation from a non-perturbed setting to two alternative scenarios to quantify the impact of changes in market prices and policy incentives on crop yields, soil health, and nutrients for the year 2014. The first scenario is based on historic occurrences where the market prices of the major crops (maize, soybeans, and wheat) are increased, resulting in higher net returns to major crop productions. In contrast, the second scenario evaluates the

impact of incentivizing grassland with a subsidy, as an illustration of the U.S. Conservation Reserve Program, with all prices and costs remaining the same as in the non-perturbed.

Major crops intensification scenario (market price change)

The first alternative scenario corresponds to an intensive crop production related to either agricultural market shocks or energy policies to expand biofuel production. The scenario assumes a 20% increase compared to the non-perturbed in maize, soybean, and wheat market prices likely resulting from higher demands for these major crops in the region. As shown in Table 1, the prices under the intensive cropping scenario fall well within the range of the most recent price surge during 2012-2013 (USDA NASS). Similarly, the economics land use model used the 2013 simulated yields for all study crops in the prairie pothole region of North Dakota to calculate the net return of each potential land use alternative under the scenario prices and transformed to the land use probability for each crop. We then implemented the ALMANAC simulations for the 2014 growing season with the scenario land use probabilities (Figure 6) under all soil and weather combinations.

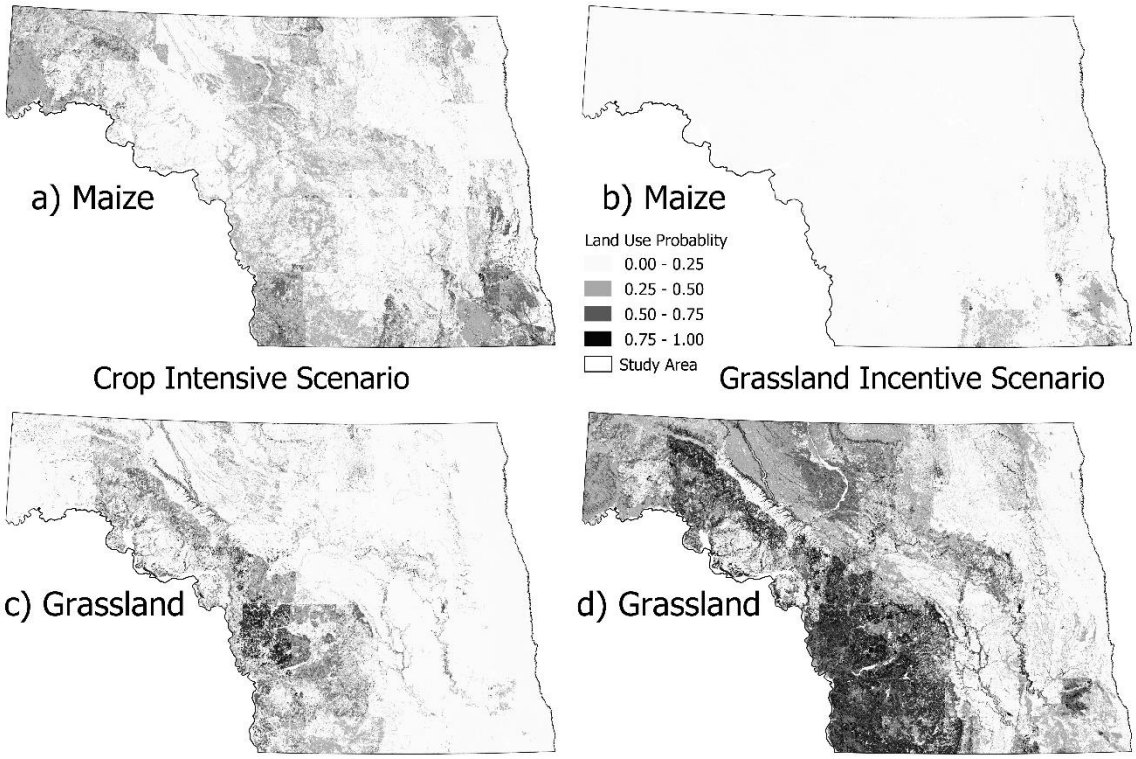

Figure 6: Probability of planting for each SSURGO soil type for maize (a and b), and grassland (c and d) using the ALM-EC model for two scenarios, crop intensive, representing the increased market prices for maize, soybean, and wheat (a and c), and grassland incentive, representing the policy change of enacting a flat per acre payment for grassland acreage (b and d). Darker gray indicates a higher probability of planting.

Grassland conservation scenario (policy change)

In contrast to the non-perturbed, where a constant net return to grassland was assigned for each land unit, the second scenario considers policies that increase subsidies or payments for ecosystems services or land rents for conservation easements such as USDA Conservation Reserve Program (CRP) land and wetland. It is assumed that an incentive of $40 per acre was added to the net return of grassland for encouraging cropland conversion to grass/pasture land with forestland remaining constant. The $40/acre rate of compensation was a midpoint of CRP rental payment ranged from $30/acre to $50/acre reported by the USDA NASS database over the past 10 years among all ND counties. With the additional $40 per acre added to the net return of grass/pasture land, it is expected that

less productive lands are more likely to remain or convert to grass/pasture use due to increased grassland profitability. To highlight the effect of this conservation effort, all other crop prices stayed at the baseline level. This generated a unique probability of planting dataset which was then fed back into the ALMANAC model to simulate the resulting yields and soil health in 2014 growing season.

3.4 Results and discussions

3.4.1 Comparison of historically-based and economics model-based crop acreages

We compared the projected land use through the ALM-EC model to both the CDL derived land use probabilities and the NASS statistics. Land use areas are aggregated to the agricultural districts and the results of which are shown in Table 2 for each of the three major crop types in the area. Note that as mentioned in section 3.3.4, yields for the seven crop types were simulated for all soil types and are used as inputs for the economics model. However, a total of 47 crop types are found through the CDL layer over the study region. We assume the land cover that is not modeled by this study, such as non-agricultural land cover and minor crops, stay constant over the study timeframe, and their acreages are removed from the analysis. However, this does not include

non-crop land uses that are accounted for by the economic model, such as grassland and forest.

Table 2: Table of 2014 land uses area in ha for each agricultural district as calculated by the ALM-EC model, the CDL derived 2010-2013 mean planted area, and the estimated planting percentages for 2014 as reported by the CDL.

		Land Use Area (ha)			Error	
Crop	Ag. District	ALM-EC Model	2010-13 CDL	2014 CDL	ALM-EC	2010-13 CDL
Maize	Central	113,948	126,410	164,308	-31%	-23%
	East Central	115,262	266,064	227,441	-49%	17%
	North Central	151,309	50,575	65,966	129%	-23%
	Northeast	87,690	89,282	84,987	3%	5%
	Northwest	145,919	11,980	17,402	739%	-31%
	Southeast	317,059	341,253	341,567	-7%	0%
	Other	65,407	75,676	88,571	-26%	-15%
Maize Total		**996,594**	**961,239**	**990,243**	**1%**	**-3%**
Soybean	Central	460,482	336,395	443,551	4%	-24%
	East Central	805,314	536,348	577,507	39%	-7%
	North Central	170,726	115,360	224,204	-24%	-49%
	Northeast	400,553	273,257	390,536	3%	-30%
	Northwest	56,628	14,489	75,679	-25%	-81%
	Southeast	497,438	440,948	564,152	-12%	-22%
	Other	20,807	31,077	79,692	-74%	-61%
Soybean Total		**2,411,948**	**1,747,875**	**2,355,320**	**2%**	**-26%**
Wheat	Central	74,443	199,669	192,233	-61%	4%
	East Central	27,599	158,064	137,598	-80%	15%
	North Central	250,607	271,498	271,554	-8%	0%
	Northeast	591,991	517,693	534,204	11%	-3%
	Northwest	603,268	230,651	418,865	44%	-45%
	Southeast	40,801	110,151	111,293	-63%	-1%
	Other	242,915	192,209	194,135	25%	-1%
Wheat Total		**1,831,624**	**1,679,935**	**1,859,880**	**-2%**	**-10%**

The historic crop hectares were derived from the historical dataset given by the CDL for the years 2010-2013 to determine the probability of each crop within each soil type within that timeframe. Finally, the 2014 data was derived from the 2014 version of the same CDL dataset to compare directly with the projected hectares from both the historical and the economics land use model. Note that the CDL measures area for each crop independently which accounts for the impacts of both non-cropland and crops not covered in this study, as such the CDL areas do not need an adjustment to account for non-study land uses.

As shown in Table 2, maize was closely predicted by both the ALM-EC model as well as the historic mean. However, while the ALM-EC model produced a more accurate estimation of the total 2014 planted area, the 2010-2013 CDL mean was able to more accurately project within most agricultural districts based on goodness of fit (estimated using R Project for Statistical Computing). Overall the ALM-EC model performs well for the region but does not factor in some of the limitations of planting such a resource intensive crop in regions not historically seen. Land use dedicated to soybean for the whole area is well projected by the ALM-EC model, surpassing the accuracy of the 2010-2013 CDL means both overall and within the majority of agricultural districts. In contrast to maize, the ALM-EC model's expansion of crops into regions not historically planted contributed to an increase in accuracy, judged by goodness of fit, within these divisions. Finally, ALM-EC's projection of spring wheat land use area is a noticeable improvement over the 2010-2013 CDL means for the whole area, but similar to maize does poorer than the 2010-2013 CDL means within most of the individual agricultural districts. Unlike maize this inaccuracy appears to be centered around the regions that experienced little to

no change in planted area from the past 3 years. As illustrated, by configuring economic factors to current market and policy conditions the ALM-EC model can reproduce the landscape of a specified year with similar, and in some cases better (e.g. Soybean), accuracy compared with using a historic mean approach.

3.4.2 Impacts of the economics model-based simulation to crop yields

With the use of agricultural land use probability from the economics model as described in Section 3.4.1 as inputs and with the methodology described in section 2.4, crop yields were simulated based on the ALMANAC model for 2014 for maize, soybean and spring wheat. Figure 7 shows the range of simulated yields in counties where the crops were commonly seen to be grown, derived from NASS quickstats total planted area for 2014 (USDA-NASS 2018). The box whisker chart is used to compare simulated crop yields, contrasted with NASS records for the mean, countywide yield for the same year represented by the red dot.

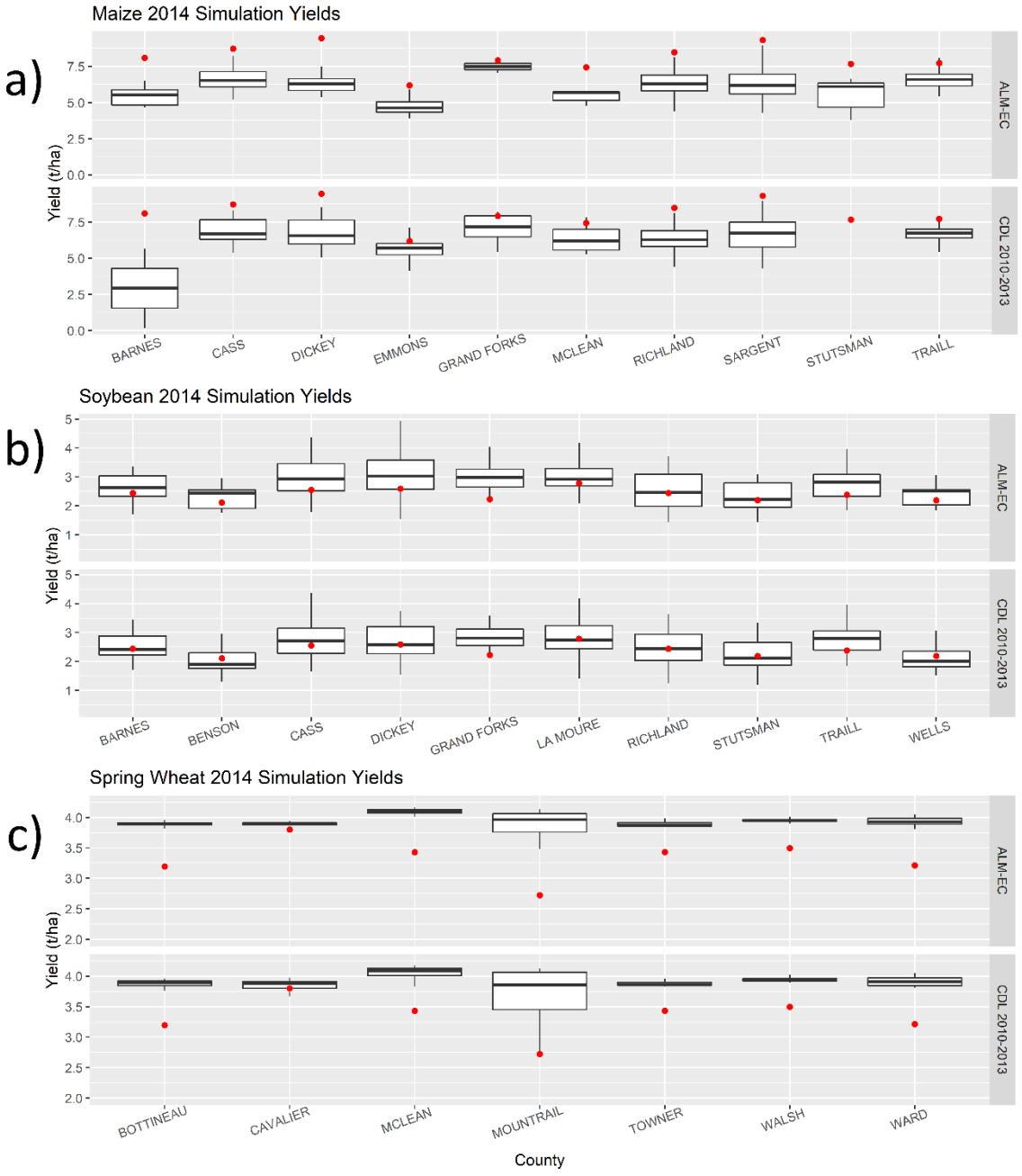

Figure 7: Simulated yields for simulations including the ALM-EC model, or the ALM-EC scenario (upper) and the static area CDL scenario (lower) scenarios of maize (a), soybean (b), and spring wheat (c) using box whisker for simulated yields, mean yields for each county as reported by NASS represented as a red dot. Simulation medians represented by the solid line.

Here the ALM-EC scenario refers to the ALM-EC model configured to replicate conditions as found in 2014, whereas the CDL scenario refers to the static-area based

standalone ALMANAC simulations based on land use mean from 2010-2013 as determined from the CDL. Simulated yields of spring wheat and soybean under the ALM-EC system remain relatively unchanged compared to the results of the static CDL-based scenario.

In contrast, maize yields under the ALM-EC overall declined compared to the NASS reported yields. This yield underperformance is likely caused by simulated expansion of maize acreage into sub-optimal productivity soils both within individual LCCs and migration to poorer productivity LCCs resulting in lower overall yields. Ultimately this expansion to poorer productivity soils is due to limited quantities of the higher quality soils, once saturation of these higher productivity soils is reached, lower quality soils are used to fulfil the remaining demand. A factor of the overall low range seen in both scenario's maize yields may be a result of the crop moving outside of the calibration soils, as only frequently planted sites between 2001-2013 were included in calibration. This greatly limited the amount of soils used during calibration, with only the highest quality soils seeing frequent maize planting during that period, relative to soybeans and wheat. Still, other factors, such as soil profile inaccuracies, calibration method, and model processes could add to this yield discrepancy.

Using the ALM-EC model resulted in soybean yields of 2.84 $t \cdot ha^{-1}$, which is 22% higher than NASS reported for the region (2.33 $t \cdot ha^{-1}$), and wheat yields of 3.88 $t \cdot ha^{-1}$, which is 17% higher than NASS reported (3.29 $t \cdot ha^{-1}$). Similarly, simulated yields were found to be 21% and 20% higher than NASS reported for soybean and wheat respectively for crop simulations with the use of historic area. The estimated maize yields are 6.02 $t \cdot ha^{-}$

^{1} and 6.88 t·ha^{-1}, which are 25% and 14% lower than NASS reported average (7.97 t·ha^{-1}) for crop simulations with the non-perturbed and the historic scenario respectively.

While a yield underperformance is found for maize, the overall performance of crop simulations, through the use of land use probabilities generated from the economics land use model, compare reasonably well with the crop simulations using CDL based land use acreage. Distinct from the standalone crop simulations, the ALM-EC modeling system treats economic factors, such as policy and market changes, as fully incorporated variables, which enables the feasibility of studying the influence/sensitivity of market and policy on crop simulations.

3.4.3 Proof-of-concept study of the impact of market and policy on crop simulations

Using the developed ALM-EC system, as a proof-of-concept study, we have perturbed the market and policy conditions for two scenarios as explained in section 3.3.5, for a total of three competing simulations: a non-perturbed simulation with a "business-as-usual" assumption to represent the conditions of the 2014 growing season with no changes (non-perturbed); a perturbation in market forces where the price of maize, soybean, and wheat are increased (crop intensive); a perturbation in policy changes where grassland is incentivized through a flat per acre payment (grass incentive). The corresponding changes to crop acreage, crop yields and soil conditions from the three simulations are documented as below.

Impacts on Crop planted areas

With the increase in demand for major crops (maize, soybean and spring wheat) the crop intensification scenario found a significant increase in simulated planted area of two of the crops, with maize experiencing the largest impact, a 157% increase in planting area, while soybean experienced gains of 49%, comparing with the non-perturbed case (Figure 8). The other two major land uses fell with grasslands experiencing a 47% decrease and wheat seeing a drop of 24% over non-perturbed planted area. Notably, while spring wheat's price was increased in this scenario, and expanded acreages in the western half of the study area, it lost ground to maize and soybeans in the eastern half where the profits, and with it increased acreage, from those crops overwhelmed the potential spring wheat acreage gains.

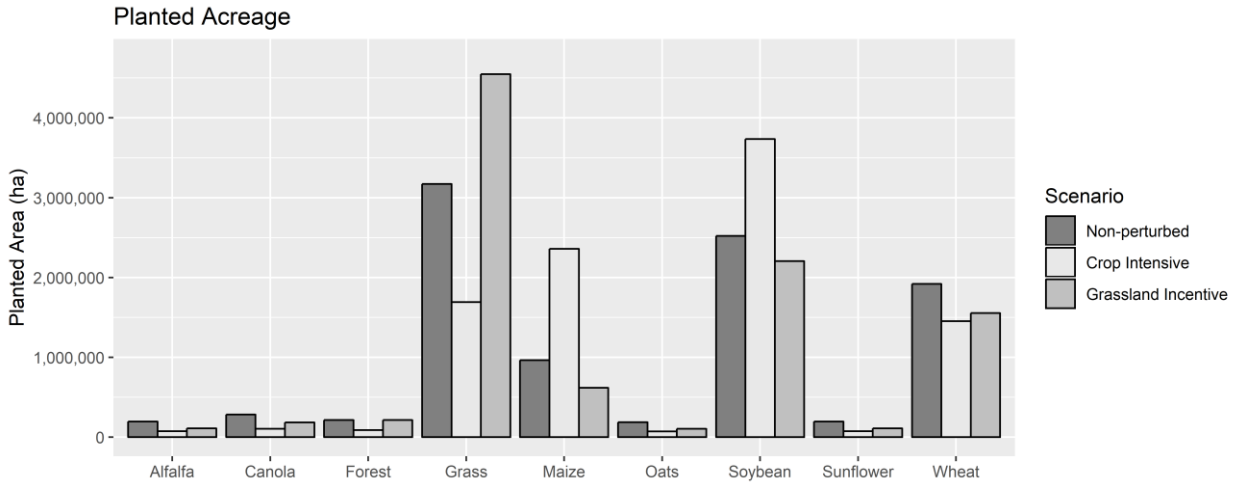

Figure 8: Total planted area of each crop under the non-perturbed scenario, crop intensification scenario, and the grassland incentive scenarios.

Conversely, with the incentivized grassland production the planted area of the major crops: maize, soybean and wheat fell by 35%, 12%, and 19% respectively while grasslands increased planting by 43% with respect to non-perturbed grassland area. In all scenarios the minor crops, such as sunflowers or oats, dropped in planting area by 34% to 71% depending on the crop and scenario as both the major crops and grasslands utilized areas previously containing these crops during their respective scenarios.

As planting incentives were modified, crops began migrating from their original state to cover LCCs not seen in the non-perturbed study. This can be demonstrated in Figure 9 where during the crop price intensive scenario a majority of the additional maize acreage went onto the higher potential productivity soils in LCC 2 (Figure 9), but also increased acreage on the less favorable LCCs. Similarly, when grassland is incentivized the majority of the acreage growth can be found at more productive LCCs, albeit at the simulated lower yielding soils within this class, while the less productive LCCs saw a smaller overall acreage growth.

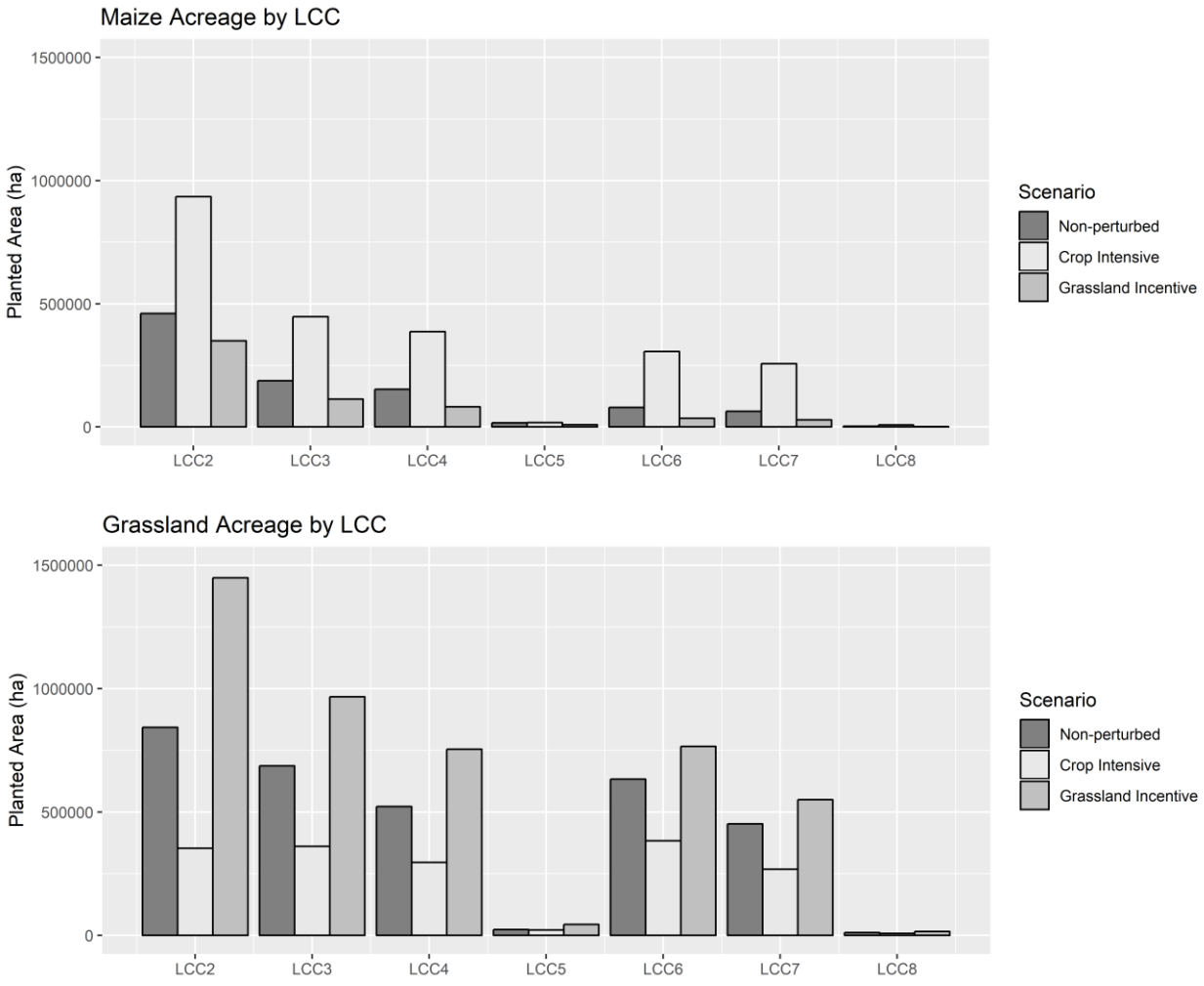

Figure 9: Total study-wide planted hectares of maize (a) and grasslands (b) by non-irrigated land capability class for the non-perturbed as well as the crop intensification and grassland incentive scenarios in hectares.

Impacts on Yields and production

Due to the expansion of acreages and prioritization of more productive soils for the more profitable crops, a similar change in yields based on the scenarios is observed (Figure 10). In the crop intensive scenario, maize, soybean, and wheat are given priority over the other four crops, resulting in expansion of those crops into less productive soils, and subsequently an overall drop in yields for maize as it's moved into the lower potential LCC (Figure 11), but not as large of an impact as its largest increase in area is found on the higher productive LCC2. Minimal impact, leading to non-significant differences in total mean yield, are found on the soybean yield as the yield drop-off from LCC2 to lower LCCs is less severe than maize during this study year and its increased area on higher productivity soil helps offset these already smaller losses. Similarly, no major change is found for spring wheat yield. Mean yield drops are found in all four of the remaining crops as those crops are forced to more marginal lands by the encroaching maize and soybean, suppressing their yields.

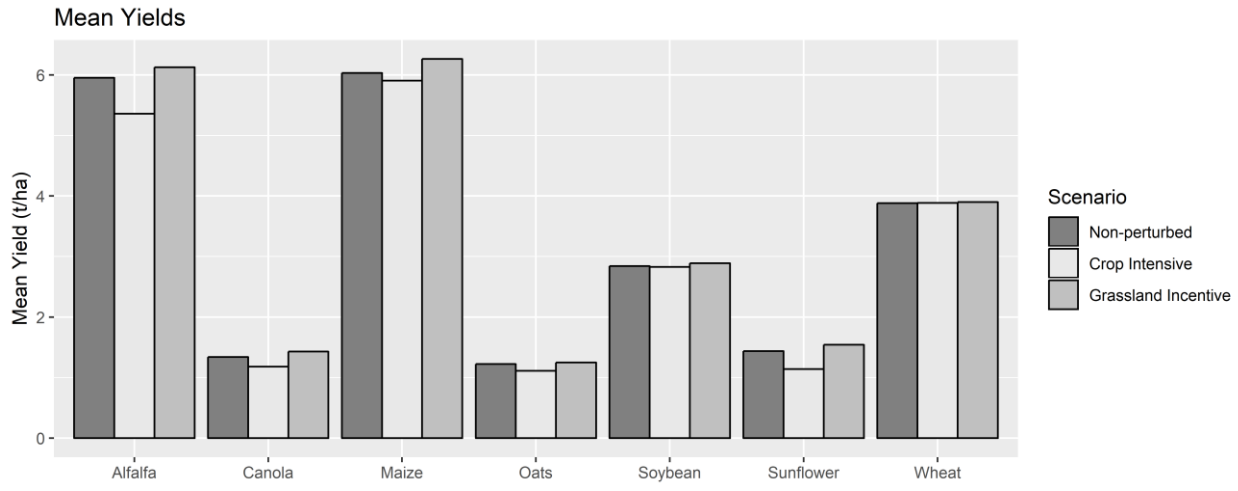

Figure 10: Mean study-wide yields for the study crops under non-perturbed, crop intensive, and grass incentive scenarios.

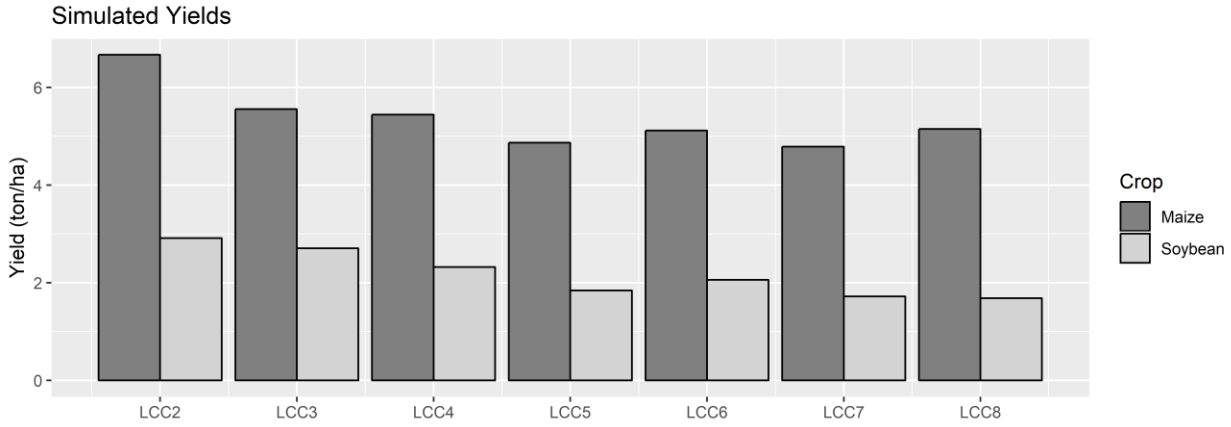

Figure 11: Mean study-wide yields for maize and soybeans crops under non-perturbed scenario grouped by LCC

However, in the grassland incentivized scenario, a grassland conversion on less productive lands limiting crop expansion is observed. As a result of this land-use change, a slight boost in mean yields are found for all crops. This is because although decreases in acreage relative to non-perturbed are observed for those crops, the poorly producing lands are removed from farming which ultimately achieves the typical conservation goal for this type of environmental policy.

Additionally, we combine the changes in acreage with the changes in yields to calculate the total production change of each study crop in the study (Figure 12). In the crop intensification study, in the study region total production increases of 152% and 48% total are found for maize and soybean respectively. This is slightly less of than the amount expected from the overall increase in acreage due to the yield impacts of planting on lower productivity soils. In contrast, the decreased in yields of alfalfa, canola, oats, and sunflowers amplifies the impact of the decrease in acreage which causes the total production of each to drop more significantly than otherwise expected when looking at each impact individually. However, for the grassland incentivized scenario, the increase in yields of the non-priority crops caused a weakening of the impacts of the acreage decrease,

where while overall production is down, the increased yields help to mitigate the loss of production relative to acreage losses. The end result being production of each crop falling relative to non-perturbed, but not to the extent we would expect if the new grassland acreage was equally distributed on all soils regardless of productivity.

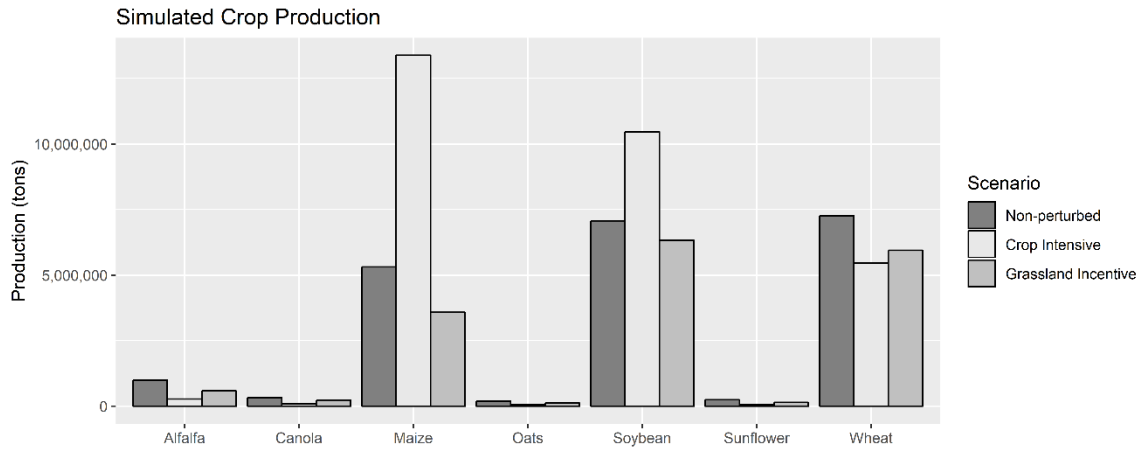

Figure 12: Total study-wide production under non-perturbed and each scenario.

3.4.4 Additional details of future scientific potential for the coupled model

The ALMANAC model also provides outputs from each soil layer about changes in soil carbon, nitrogen, and other key components in soil health. However, due to the short-term nature of this study these results for the combined ALM-EC model were generated but not formally included in this study. Given a longer-term study and sufficient calibration data the impacts of the annually changing land use could be measured with this system, which could provide valuable insight into long term changes and their subsequent impact to overall soil health. Furthermore, due to the single year status of the simulations impacts from rotations were not analyzed within this study. These effects can be handled by the ALM-EC model via an assumption that all crop movement will follow standardized rotations predefined prior to the simulations, creating best-fits for each possible

combination, or through a statistical likelihood approach using historic records of representative acreage within the area to project future probabilities.

3.5 Conclusions and discussions

In this concept demonstration study, a crop and economics modeling system is developed through a two-way linkage of an economics land use model with the ALMANAC model for crop simulation. The designed goal for this system is to include economic factors as fully incorporated variables in crop simulations, allowing the study of sensitivity of crop simulations with respect to market and policy changes. We have demonstrated and tested the concept for the 2013-2014 season over North Dakota and inter-compared the new crop simulation concept with a static approach which uses historical acreage data to simulate seven crops common to North Dakota. This study finds:

(1) For the 2014 study period, dynamic crop acreages can be generated using the ALM-EC system while producing similar performances in acreage and yields in comparison to a static, standalone crop simulation.

(2) Comparing to the non-perturbed case in market and policy conditions, crop simulations under a crop intensified scenario where market prices for major crops are increased, the model introduces increases in acreages for the in-demand crops on the remaining higher productivity soils, both of these factors impact other competing crops negatively leading to decreases in both acreage and yields.

(3) Repeating this study with a policy change that favored grasslands we observed a grassland expansion due to an increase in its net return. However, unlike the market price study we found that the grassland conversion was mainly focused on the lower productivity soils. As a result of the removal of these lower productivity soils from the available planting pool for the minor crops, yields increased and slightly offset the overall loss of acreage for these crops.

Note that in the newly developed ALM-EC, land use probabilities are derived from the economics land use model. This allows us to examine the impacts of likely crop locations and the effect of specific soil types on yields even when lacking accurate land use information, which may enable a more realistic long-term crop simulation and forecast.

This study suggests that with the use of two-way linkage between an economics land use model and a crop model, economic factors can be included as control variables and be further used to study the sensitivity of market drivers on crop simulations. This study serves as a foundational step towards the goal of reducing the impact of unrealistic depictions of static agricultural land use in both short- and long-term simulations by incorporating high-resolution dynamic land use, through a coupled economic model, into crop simulations. The developed system, in theory, may be used to gain an insight into the changes in agricultural practice due to policy and market changes for potential policy and decision making.

CHAPTER 4

Albedo Impacts of Changing Agricultural Practices in the United States through Space-Borne Analysis

4.1 Rationale

Surface albedo, which is the ratio of outgoing to incoming solar irradiance at a given wavelength, is a key component to a range of scientific applications such as the lower boundary conditions for the remote sensing of surface and atmospheric properties (Seidel, et al., 2012), climate studies (Zelinka, et al., 2020), and atmospheric modeling parameters (Jandaghian, et al., 2020). Changes in surface albedo over cropland occur at daily, weekly, and seasonal scales due to crop growth. Yet, while seasonal changes in surface characteristics are accounted for in some applications, these daily or weekly changes in surface albedo are less frequently considered. Still, significant variation in both timing and range of surface albedo are observed for crops from initial planting to harvest stages, which can be noted simply by the significant variation of color hues displayed over croplands during maturation (e.g. Figure 13).

Figure 13: National Agriculture Imagery Program (NAIP) visible aircraft image of farmland in Cass County, ND on 08/19/2017 (USDA-FSA, 2019). The spring wheat field has matured before the three neighboring crops producing a higher surface albedo relative to the nearby fields. Note that although a visible image is shown, surface albedo over crop land also drastically changes as a function of wavelength.

To maintain soil health, to suppress pests and disease, and thereby maximize long-term productivity and profits, farmers routinely use cyclical crop rotations at a given location such that the crop planted (along with pest and disease pressure) varies on an annual to semi-annual basis. As a result of historically routine crop rotation practices and new rotational practices related to market conditions and US agricultural policy, surface

characteristics of cropland derived from prior years may not be representative of current or future conditions. Additionally, in-season surface albedo variation caused by both crop rotation and crop growth throughout the growing season are a non-negligible issue due to the large spatial extent of agricultural lands. For example, in the continental United States in 2019, farmland covered a total of 895 million acres (USDA - NASS, 2020) (47.4% of total land area), with 308 million acres of that in harvestable crops (USDA - NASS, 2020) (16.3% of total land area).

Several approaches have been attempted to account for changes in surface albedo resulting from plant growth. For example, Houldcroft et al. (2009) defined albedo parameters for five vegetation categories (and the corresponding bare soil) worldwide using MODIS derived albedo, calculating albedo as a combination of plant and soil reflectivity. Hsu et al. (2019) used the normalized difference vegetation index (NDVI) to account for cropland surface albedo changes for satellite aerosol retrievals, whereas Zhang et al. (2020) produced time-series forecasts for albedo of a specific area including croplands through empirical mode decompositions and neural networks. Still, for most prior research applications, surface albedo values have either been assumed to be invariant throughout the study period or estimated through the use of a static look-up table based on generic cropland albedo changes.

In this study, through collocation of the crop land data layer (CDL) with the Moderate Resolution Imaging Spectroradiometer (MODIS) bidirectional reflectance distribution function data, the variations of surface albedo during the crop growing season are studied for 55 crop types, seven spectral channels (ranging from visible to shortwave infrared), and nine different plant hardiness zones across the United States.

From this collocation a database was created for applications that require accurate estimations of daily surface albedo changes over cropland. Using this database this study tackles three overarching questions: 1) what are the impacts of crop variations in growth and reflectivity on spectral surface albedo, and are these impacts consistent on a per-wavelength, crop type, day of year, and hardiness zone basis?; 2) can cropland NDVI values be used as a proxy for the variations in surface albedo over cropland caused by plant growth?; and, 3) are the broadband albedo changes due to crop growth significant for climate applications?

4.2 Methodology

The changes in spectral albedo as a function of crop growth were studied over the continental United States during a four-year period (2015-2018). Four datasets, including USDA-NASS's CDLs, USDA's Plant Hardiness Zones database, and MODIS derived albedo and reflectivity were collocated and used for the study. The study's region and timing were chosen due to the availability and consistency of the CDL data, which allows for identification and inclusion of a large number of homogeneous single-crop locations overlaying MODIS pixels.

4.2.1 Collocation of CDL and MODIS data

To calculate the seasonal change in albedo of specific crops requires collocation of the CDL with the MODIS gridded BRDF dataset. Figure 14 shows the flow chart of the collocation steps. Colocation was performed by locating the center of each BRDF grid pixel and locating the corresponding CDL pixel found at that location. Due to the resolution difference (30m CDL vs 500m MODIS) an appropriate buffer area around the

BRDF pixels was selected extending beyond the MODIS pixel. All CDL pixels within the area + buffer were checked for single crop homogeneity, with the overlaying BRDF pixel discarded from the study if any pixel within the area + buffer was classified as a different crop or land cover from the central CDL pixel. As a result, all MODIS BRDF pixels used in this study are considered to be comprised entirely of a single crop as interpreted by the CDL. The remaining, single crop, BRDF pixels were grouped by crop and HZ, and the mean albedo, standard deviation of that albedo, and raw pixel counts for each day of the year calculated from the merged four-year BRDF dataset.

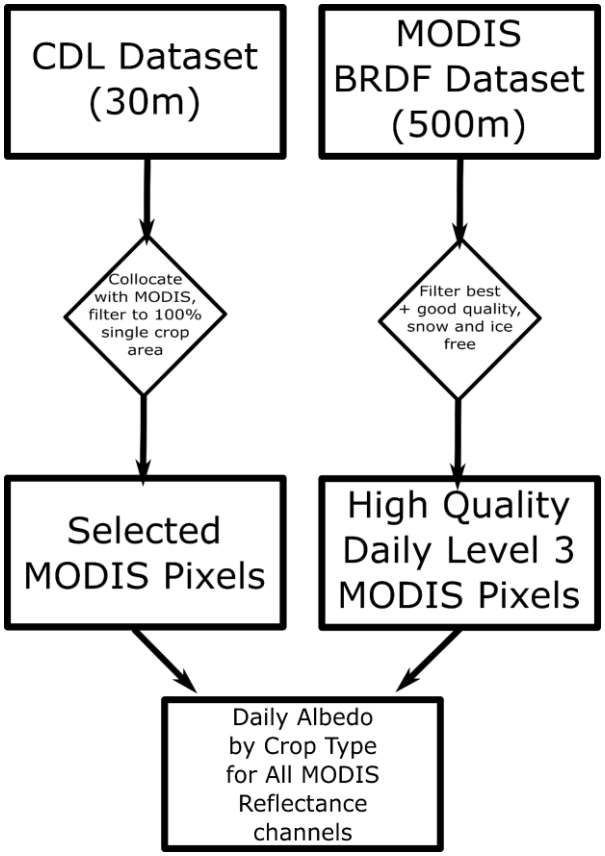

Figure 14: Flowchart for pixel selection and calculation

As the result of colocation, 20,345,626 MODIS pixels across all land cover types were found that matched the criteria over the four year period (roughly 5,000,000 per

year); which are shown below grouped by land cover type (figure 15a), by HZ (figure 15b), and the selected points of several common US crops (figure 16) for reference. Using these reference points both mean and standard deviation of albedo for each crop, HZ, Julian day, wavelength (including broadband), and sky-type combination were calculated. Additionally, means and standard deviations of the albedo for the remaining homogeneous, but non-cropland, land covers were derived and included in the final database, but not focused upon in this study.

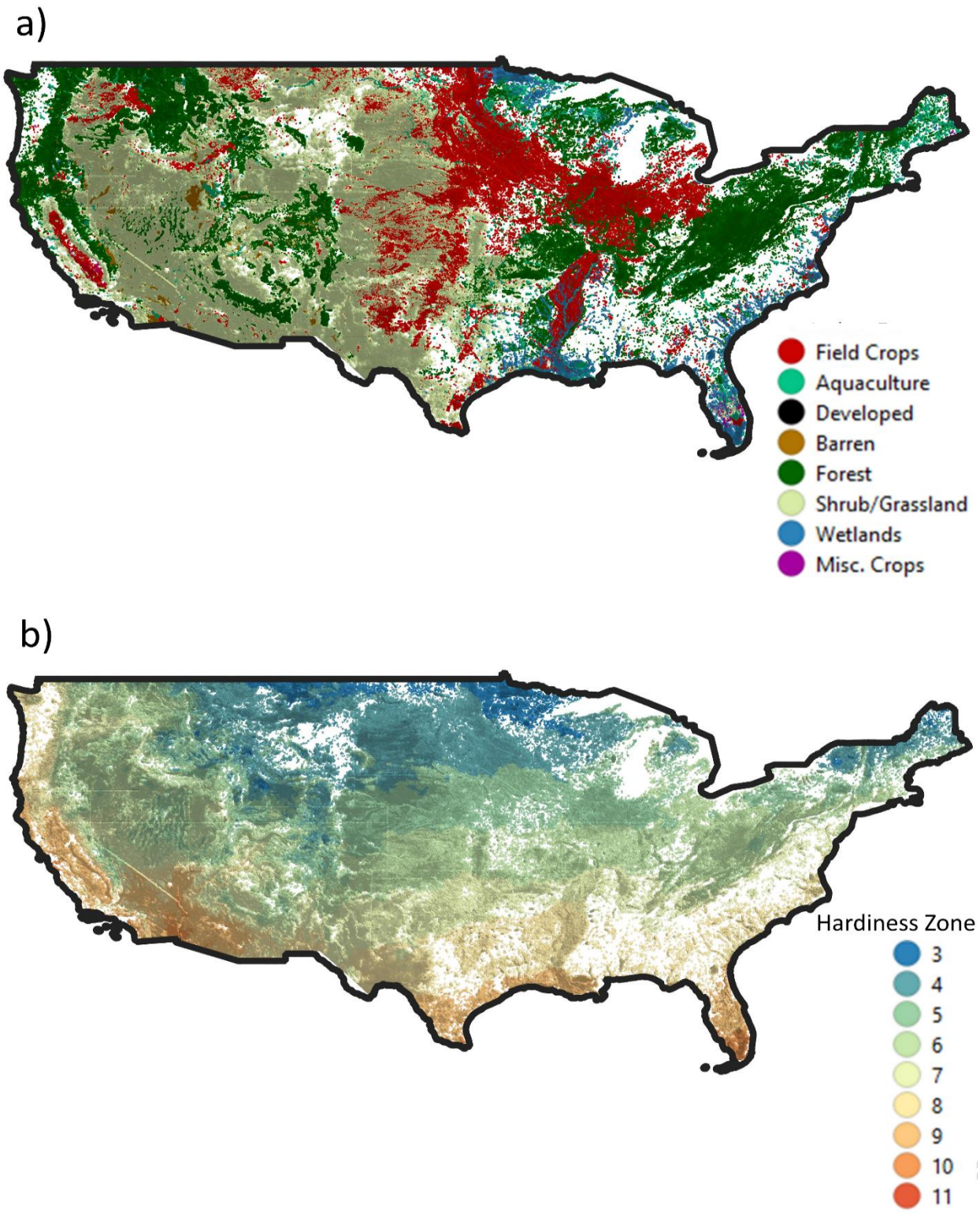

Figure 15: Selected Points by a) landcover type and b) HZ as defined by USDA

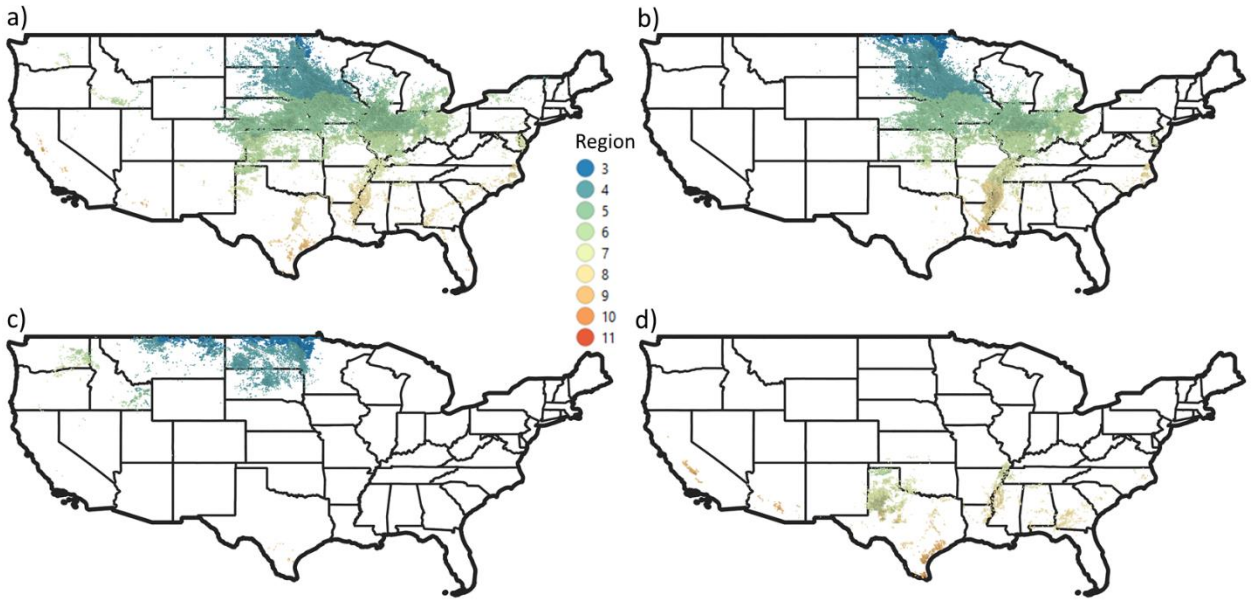

Figure 16: Locations of selected points of in this study, colored by plant hardiness zone for four example crops of (a) maize, (b) soybeans, (c) spring wheat, and (d) cotton

4.3 Results and discussions

The black- and white-sky albedo are studied as functions of crop type and growing period over the four-year study period in this section. Again, the goal of this effort is to construct a database providing for the fast derivation of albedo changes over cropland through the growing period, examples of the seasonal variations of which are shown in Table 3. The database includes the mean and standard deviation of the albedo, at each MODIS visible and near IR wavelength, including broadband, for both sky types, for every Julian day of the year for a total of 55 crop types. Also included in the dataset are ancillary parameters such as valid MODIS pixel counts. The dataset is included as a supplement for this study. Examples of the albedo changes as a function of growing days at visible, near IR, and SWIR channels, as well as the broadband SW spectrum, are shown in the following sections for maize, soybean, spring wheat and cotton. These are four of the most commonly planted spring to fall cycle crops in the US, and are shown as

examples (Maize with 91.7mil acres, soybeans with 80.0mil acres, spring wheat with 12.4mil acres, and cotton with 13.7mil acres (USDA - NASS, 2020).

Table 3:Mean albedo range of common US crops (HZs 3-6) for the growing season (June-August) by MODIS channel wavelength (nm)

		BLACK SKY ALBEDO							WHITE SKY ALBEDO						
CROP	**HZ**	**470nm**	**555nm**	**645nm**	**860nm**	**1240nm**	**1640nm**	**2130nm**	**470nm**	**555nm**	**645nm**	**860nm**	**1240nm**	**1640nm**	**2130nm**
ALFALFA	4	0.034-0.062	0.076-0.105	0.062-0.116	0.346-0.397	0.340-0.375	0.234-0.303	0.125-0.179	0.038-0.070	0.087-0.118	0.068-0.128	0.381-0.445	0.381-0.406	0.258-0.326	0.134-0.188
ALFALFA	5	0.040-0.052	0.080-0.094	0.074-0.097	0.367-0.386	0.337-0.366	0.239-0.274	0.124-0.153	0.045-0.058	0.092-0.105	0.083-0.108	0.406-0.436	0.377-0.398	0.263-0.298	0.132-0.164
ALFALFA	6	0.045-0.057	0.085-0.101	0.083-0.106	0.350-0.376	0.328-0.359	0.240-0.276	0.144-0.163	0.051-0.064	0.098-0.113	0.093-0.118	0.394-0.413	0.364-0.389	0.266-0.301	0.156-0.176
BARLEY	4	0.032-0.088	0.075-0.140	0.058-0.183	0.321-0.404	0.322-0.391	0.215-0.344	0.108-0.204	0.035-0.093	0.088-0.151	0.064-0.195	0.371-0.466	0.363-0.419	0.241-0.367	0.115-0.213
BARLEY	5	0.026-0.089	0.068-0.144	0.048-0.186	0.363-0.457	0.326-0.399	0.178-0.338	0.078-0.208	0.030-0.097	0.086-0.156	0.057-0.198	0.394-0.524	0.382-0.437	0.210-0.360	0.087-0.217
CANOLA	3	0.026-0.073	0.077-0.120	0.051-0.147	0.218-0.483	0.285-0.368	0.161-0.311	0.064-0.238	0.028-0.077	0.086-0.129	0.056-0.156	0.250-0.533	0.316-0.393	0.180-0.335	0.071-0.255
DRY BEANS	4	0.033-0.067	0.075-0.112	0.063-0.133	0.251-0.390	0.304-0.366	0.225-0.309	0.109-0.221	0.037-0.071	0.085-0.120	0.069-0.142	0.286-0.441	0.336-0.402	0.246-0.326	0.116-0.229
DURUM WHEAT	3	0.030-0.082	0.073-0.130	0.057-0.174	0.253-0.377	0.316-0.386	0.218-0.361	0.104-0.225	0.031-0.089	0.083-0.142	0.060-0.190	0.286-0.440	0.345-0.417	0.236-0.385	0.107-0.230
FALLOW/IDLE	3	0.053-0.079	0.091-0.119	0.102-0.151	0.250-0.273	0.306-0.331	0.291-0.359	0.187-0.248	0.055-0.084	0.098-0.129	0.107-0.161	0.272-0.319	0.340-0.358	0.308-0.378	0.191-0.257
FALLOW/IDLE	4	0.053-0.082	0.092-0.124	0.102-0.158	0.251-0.279	0.310-0.344	0.292-0.368	0.183-0.247	0.055-0.089	0.099-0.134	0.106-0.170	0.272-0.325	0.346-0.370	0.308-0.386	0.187-0.255
FALLOW/IDLE	5	0.061-0.073	0.104-0.115	0.122-0.144	0.272-0.281	0.319-0.338	0.308-0.346	0.210-0.241	0.063-0.079	0.111-0.127	0.126-0.156	0.299-0.325	0.351-0.376	0.323-0.364	0.213-0.247
FALLOW/IDLE	6	0.062-0.082	0.104-0.128	0.125-0.165	0.275-0.294	0.312-0.351	0.299-0.346	0.211-0.245	0.065-0.087	0.113-0.138	0.131-0.177	0.317-0.328	0.348-0.381	0.318-0.370	0.217-0.256
MAIZE	3	0.017-0.039	0.057-0.072	0.033-0.071	0.210-0.439	0.265-0.362	0.191-0.267	0.073-0.203	0.021-0.042	0.068-0.081	0.039-0.077	0.245-0.488	0.300-0.402	0.214-0.293	0.081-0.215
MAIZE	4	0.018-0.046	0.053-0.078	0.034-0.089	0.213-0.452	0.283-0.366	0.199-0.309	0.076-0.241	0.021-0.050	0.063-0.086	0.039-0.096	0.243-0.500	0.314-0.407	0.226-0.335	0.085-0.257
MAIZE	5	0.018-0.052	0.050-0.089	0.032-0.103	0.258-0.455	0.321-0.363	0.199-0.326	0.074-0.245	0.021-0.055	0.059-0.096	0.038-0.108	0.290-0.506	0.349-0.405	0.226-0.350	0.083-0.256
MAIZE	6	0.022-0.055	0.056-0.096	0.042-0.108	0.295-0.422	0.333-0.353	0.211-0.313	0.090-0.220	0.026-0.058	0.066-0.103	0.048-0.114	0.335-0.468	0.363-0.395	0.238-0.331	0.099-0.228
HAY/NON ALFALFA	4	0.031-0.049	0.070-0.086	0.058-0.093	0.289-0.319	0.310-0.349	0.220-0.297	0.108-0.161	0.035-0.055	0.085-0.098	0.065-0.104	0.331-0.388	0.363-0.388	0.249-0.321	0.116-0.172
HAY/NON ALFALFA	5	0.030-0.053	0.069-0.094	0.056-0.103	0.321-0.353	0.306-0.371	0.205-0.296	0.103-0.160	0.035-0.060	0.086-0.107	0.065-0.117	0.362-0.421	0.364-0.408	0.237-0.322	0.114-0.173
SORGHUM	5	0.047-0.079	0.089-0.124	0.088-0.153	0.283-0.350	0.332-0.354	0.284-0.353	0.170-0.255	0.052-0.083	0.100-0.133	0.098-0.161	0.317-0.383	0.359-0.388	0.309-0.374	0.181-0.259
SORGHUM	6	0.042-0.071	0.082-0.113	0.081-0.141	0.276-0.354	0.324-0.349	0.273-0.344	0.160-0.238	0.048-0.075	0.094-0.122	0.091-0.150	0.312-0.387	0.354-0.382	0.299-0.362	0.174-0.243
SOYBEANS	3	0.020-0.041	0.061-0.083	0.040-0.079	0.197-0.413	0.250-0.363	0.201-0.256	0.081-0.191	0.022-0.044	0.071-0.091	0.044-0.086	0.226-0.460	0.281-0.398	0.223-0.278	0.087-0.203
SOYBEANS	4	0.019-0.048	0.055-0.080	0.037-0.094	0.216-0.456	0.286-0.376	0.214-0.310	0.084-0.239	0.022-0.051	0.063-0.087	0.042-0.099	0.245-0.498	0.316-0.413	0.235-0.335	0.091-0.254
SOYBEANS	5	0.015-0.053	0.046-0.089	0.027-0.104	0.253-0.488	0.319-0.384	0.207-0.328	0.073-0.246	0.018-0.054	0.054-0.094	0.031-0.108	0.284-0.529	0.348-0.419	0.229-0.352	0.081-0.257
SOYBEANS	6	0.017-0.054	0.049-0.094	0.031-0.105	0.284-0.461	0.329-0.374	0.211-0.312	0.079-0.221	0.019-0.056	0.056-0.100	0.035-0.110	0.325-0.499	0.361-0.409	0.231-0.333	0.086-0.231
SPRING WHEAT	3	0.024-0.063	0.065-0.107	0.047-0.130	0.230-0.409	0.274-0.360	0.195-0.304	0.088-0.190	0.027-0.067	0.076-0.116	0.052-0.140	0.267-0.470	0.308-0.398	0.220-0.322	0.096-0.200
SPRING WHEAT	4	0.028-0.060	0.067-0.103	0.054-0.126	0.281-0.375	0.316-0.361	0.219-0.318	0.107-0.188	0.031-0.065	0.078-0.114	0.059-0.138	0.324-0.438	0.351-0.402	0.242-0.335	0.112-0.196
SPRING WHEAT	6	0.032-0.100	0.069-0.159	0.063-0.222	0.338-0.368	0.309-0.413	0.208-0.371	0.114-0.225	0.035-0.106	0.081-0.171	0.069-0.238	0.393-0.434	0.358-0.440	0.231-0.395	0.119-0.236
SUGARBEETS	4	0.019-0.041	0.062-0.079	0.037-0.073	0.202-0.453	0.255-0.359	0.181-0.267	0.071-0.212	0.022-0.044	0.073-0.086	0.042-0.078	0.229-0.500	0.283-0.400	0.207-0.290	0.078-0.226
SUNFLOWER	4	0.034-0.055	0.077-0.093	0.068-0.110	0.275-0.387	0.333-0.357	0.244-0.329	0.118-0.211	0.038-0.056	0.087-0.099	0.077-0.114	0.316-0.429	0.364-0.401	0.268-0.345	0.126-0.215
WINTER WHEAT	3	0.039-0.084	0.077-0.128	0.073-0.167	0.279-0.334	0.309-0.353	0.240-0.364	0.144-0.236	0.042-0.089	0.088-0.138	0.078-0.180	0.303-0.402	0.355-0.383	0.257-0.385	0.147-0.245
WINTER WHEAT	4	0.034-0.080	0.072-0.125	0.065-0.165	0.278-0.353	0.314-0.364	0.230-0.359	0.128-0.216	0.038-0.085	0.085-0.136	0.071-0.178	0.304-0.418	0.366-0.398	0.254-0.378	0.135-0.224
WINTER WHEAT	5	0.044-0.067	0.084-0.109	0.084-0.139	0.275-0.331	0.315-0.344	0.250-0.330	0.160-0.211	0.049-0.076	0.099-0.127	0.094-0.158	0.306-0.389	0.361-0.389	0.279-0.351	0.170-0.220
WINTER WHEAT	6	0.045-0.066	0.084-0.110	0.092-0.142	0.293-0.306	0.305-0.352	0.249-0.332	0.162-0.221	0.052-0.071	0.101-0.119	0.106-0.152	0.329-0.365	0.353-0.384	0.279-0.356	0.173-0.232

4.3.1 Variations in cropland albedo at visible, NIR, and SWIR channels due to the crop growth cycle

Figure 17 shows the change in black-sky surface albedo as a function of the growing period for maize, soybeans, spring wheat, each in HZ4, and cotton, in HZ8, for the seven spectral MODIS channels (470nm, 555nm, 645nm, 860nm, 1240nm, 1640nm, 2130nm). The solid lines show the daily mean albedo for the seven wavelengths, while the shaded area shows the standard deviation. For illustration purposes, for maize, soybeans, and spring wheat, HZ4 is shown as the spatial overlap of these three crops is high in this HZ, allowing for direct comparison. However, as displayed in Figure 16, cotton does not appear in HZ4 and as such HZ8 is included for reference, but does not spatially overlap with the other three crops.

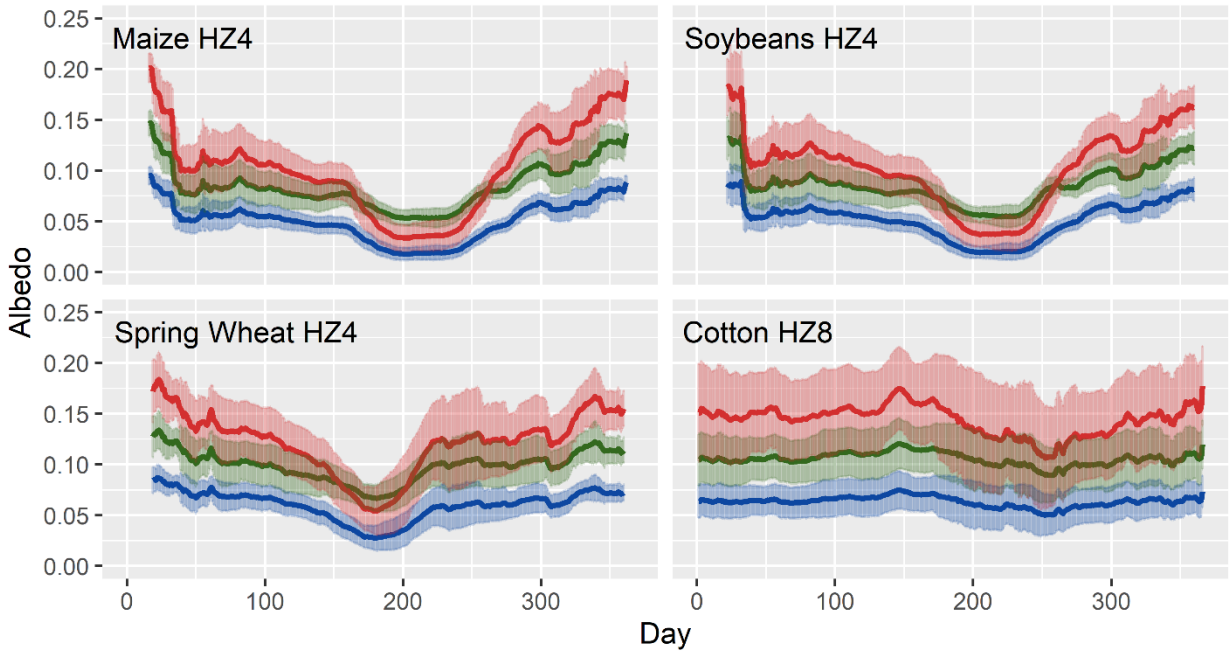

Figure 17: Visible mean (solid line) and standard deviation (bars) of BSA for four popular crops in selected HZs for 470nm (blue), 555nm (green), and 645nm (red)

As shown in Figure 17, significant changes in black-sky cropland albedos in the visible spectrum, particularly with respect to the 645 nm channel, are clearly observable. For maize fields, black-sky albedo is reduced from 0.1 at the beginning of the planting season (day 100) to a low of 0.03 during the peak growing season (day 200), representing a 70% relative decrease in black-sky albedo due to plant photosynthesis. Following the peak growth period, the black-sky albedo for maize fields increases to 0.15 to 0.2 during and after the harvesting period (after day 275), as plant photosynthesis activity diminishes during crop maturity and the crop is ultimately removed from the surface.

Near identical patterns in black-sky albedo are found for soybean fields compared with maize fields, suggesting that the two crops share similar growth and photosynthesis activity cycles. In comparison, spring wheat fields are planted earlier and mature at a much faster rate relative to maize and soybeans. Consequently, the minimum black-sky albedo is centered around late June (day 175), which is around 25 days earlier than soybean or maize fields (day 200), and the change is sustained for a much shorter period of time.

The black-sky albedo change for cotton is shown in Figure 5d for HZ8, which again is chosen as a representative zone as cotton is not commonly found in HZ4. Note that cotton fields in the United States are planted around mid-May (day 150) and harvested in late October (day 290). Still, both greater variance as well as a less significant overall change, with the black-sky albedo changing from 0.10 to 0.15, was found during the peak growing season. This lack of definable response could be a result

of the agricultural practices around cotton; as cotton has a wider range of planting and harvesting dates than maize and soybean, as well as the extended growth cycle relative to the other US annual crops based on the NASS' annual reports (USDA - NASS, 2010).

In comparison with the 645 nm spectral channel, lower black sky albedo values are found for both 470nm and 555nm channels for most crops. For example, maize field albedo values are around 0.1 for the 645nm channel during the planting period while the values are around 0.075 and 0.05 for 555nm and 470nm spectral channels, respectively. Lower relative changes between planting and peak growing seasons are also found, with a 50% decrease in black-sky albedo for the blue 470nm channel and a similarly muted 33% reduction in black-sky albedo in the green channel. This change is in line with expectations of albedo changes caused by photosynthesis activity as the majority of chlorophyll absorption is found in the red and blue areas of the spectrum, with relatively less absorption of green by chlorophyll *a* and b (Chappelle, et al., 1992).

Plants have a high albedo at the 860 nm spectral channel, and therefore, as LAI increases nearly opposite patterns from the visible channels are observed for the 860 nm channel (e.g. Figure 18). Lower black-sky albedo values on the order of 0.2 are found for maize, soybeans, and spring wheat during the planting season, while maximum mean black-sky albedo values around 0.4 are found during the peak growing season as identified in Figure 17. A similar, however more muted, increase is observed in black-sky albedo for cotton fields, with an increase from .25 to .35 from planting through peak growing season.

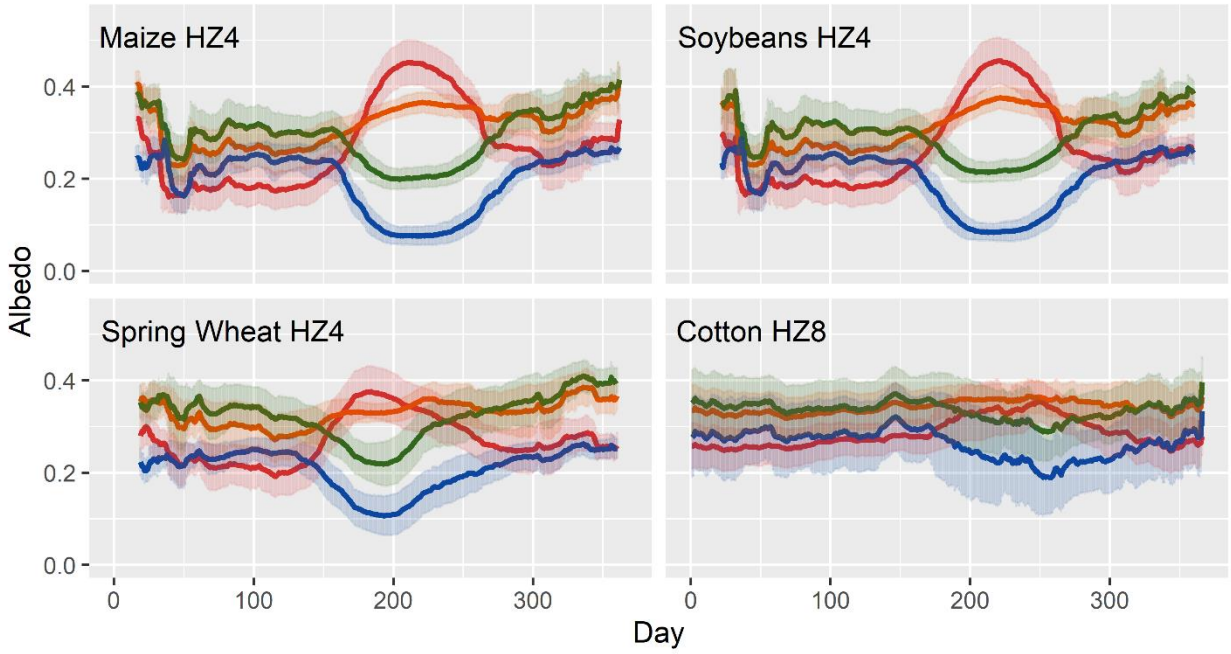

Figure 18: IR mean (solid line) and standard deviation (bars) of BSA for four popular crops in selected HZs for 860nm (red), 1240nm (orange), 1640nm (green), and 2130nm (blue)

The black-sky albedo values at 1640 nm channel behave similar to the visible channels but with a more drastic change in absolute magnitude between planting and peak growth periods. The changes in mean black-sky albedo values are from 0.3 to 0.2, 0.3 to 0.22, and 0.35 to 0.22 for maize, soybean, and spring wheat, respectively, from planting to peak growth periods. A 0.05 change in black sky albedo value from 0.35 to 0.3 is also found for cotton fields from the start of planting to peak growth. Similar patterns are also seen for the 2140 nm channel, although the absolute values are lower and the seasonal change is relatively more magnified. Surface albedo values of soil at these wavelengths is relatively high in comparison with healthy plants (Richards, 2013) which would explain the more muted response with respect to the lower wavelengths.

Distinct patterns, however, are observed for the 1240 nm spectral channels for most crops. In general, lower black-sky albedo values are found at the planting stage, with mean values of 0.25, 0.25, 0.3 and 0.33 for maize, soybean, spring wheat and cotton, respectively, that increases to approximately 0.35 for all four crop types around the peak of the growing season. Still, two local maximum black sky albedo values at day 160 and day 220 are found for spring wheat, which represents a unique pattern relative to all other channels. Also, higher black sky albedo values are also found during and after harvesting seasons that is also in contradiction to patterns as observed from other channels. Whereas similar albedo values are found for planting and harvesting seasons.

The white-sky albedo values show similar behavior as the black-sky albedos, and similar figures as Figures 17 and 18 are included in the supplement document. In summary, drastic changes in both black- and white-sky albedos are found, on the order of 30-75% for the major United States crops of maize, soybean, and spring wheat during the growth cycle for visible, NIR, and SWIR MODIS spectral channels. A similar, but lesser, change in magnitude is also found for cotton fields. Clearly, changes as significant as this in surface albedo during the plant growing period, as a function of plant type, may need to be taken into account when undertaking either remote sensing applications that require a knowledge of lower boundary conditions over cropland, or regional to global scale numerical weather/climate prediction and simulation models.

Of particular importance is the predictability of the annual cycle in cropland albedo as defined independently by each crop type. If there are significant annual inconsistencies in this cycle, on a per-crop basis, driven by external factors then the

overall benefit of applying these albedo changes to future studies decreases substantially. To investigate the annual variations and their potential impacts we separated each crop into its annual components and applied the earlier analysis over all wavelengths and albedo types independently. As observed in Figures 19 and 20 crops typically followed the same pattern of albedo changes, indicating the overall annual pattern as found from the study are recurring patterns.

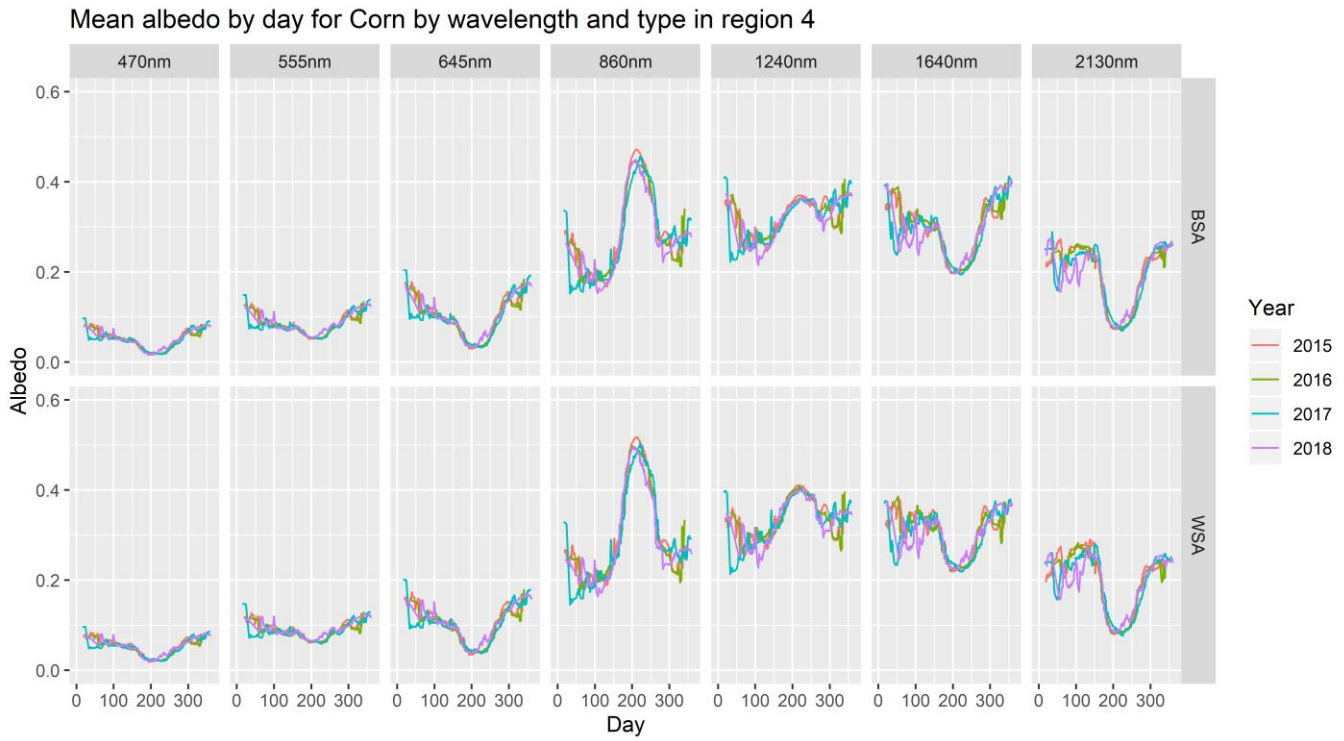

Figure 19: Year-by-year variation in albedo in HZ4 for maize for all MODIS reflectance channels, black-sky albedo (BSA) & white-sky albedo (WSA)

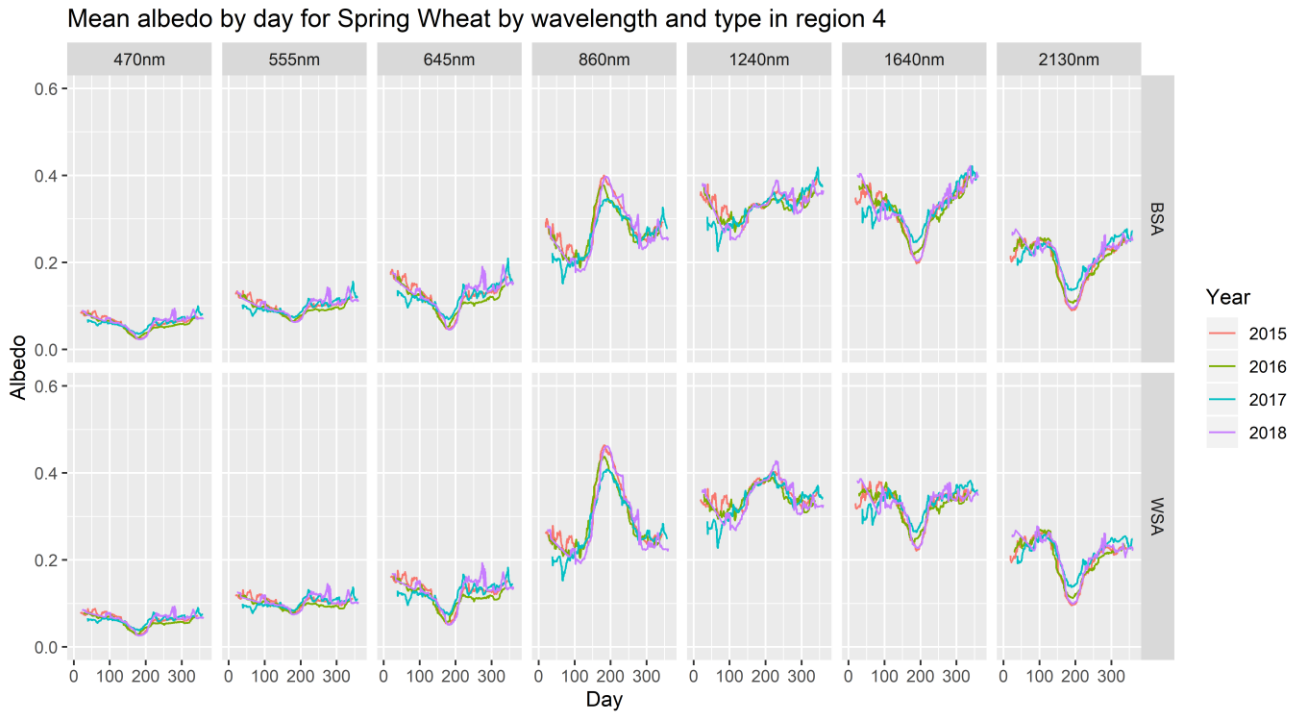

Figure 20: Year-by-year variation in albedo in HZ4 for spring wheat for all MODIS reflectance channels, black-sky albedo (BSA) & white-sky albedo (WSA)

4.3.2 Variations in broadband SW cropland albedo due to the crop growth cycle

In addition to the individual channels, the MODIS BRDF product provides broadband albedo calculations in visible (0.3-0.7nm), NIR (0.7-3.0nm), and shortwave (SW; 0.3-3.0nm) (Schaaf, et al. 2002). These have been compiled in a similar manner to the individual MODIS channels in this study with the mean and standard deviation calculated for each crop, HZ, and day of year. While the impact of crops on SW broadband albedo is less significant than the impacts on the individual channels, it is still non-negligible.

Figure 21a shows the mean broadband black-sky SW albedo for maize, and spring wheat for HZ4. Of the two crops, maize is found to have the larger seasonal swing in SW albedo, with an average SW black-sky albedo of 0.13 during the planting and early growing season with the SW black-sky albedo values changing to 0.17-19 during and after harvesting seasons. This still represents around a 30% change in broadband black-sky albedo during the crop growing period, which could have a non-negligible impact on regional climates. A similar, but lesser, change in magnitude is found for spring wheat where the SW black-sky albedo values are observed to undergo a 10-20% change during the growing period. The relative change of SW black-sky albedo, however, is observed to be only ~10% for cotton fields. Very similar patterns are found for the white sky albedo, with an approximately 30% change in SW white-sky albedo for both corn and spring wheat as shown in Figure 21b.

Figure 21: (a) Total shortwave, black-sky albedo by Julian day of maize (red) and spring wheat (blue) in HZ4. (b) Total shortwave, white-sky albedo by Julian day of maize (red) and spring wheat (blue) in HZ4.

As suggested from Figure 21, variation in broadband albedo exists for different HZs and crops due to a range of reasons including distinct planting date ranges, growing season length, and local crop varieties. For example, cropland planted in maize in HZ4 reaches the highest albedo around August 4th, whereas further south in HZ8 the peak in albedo can be found a full month earlier, around July 1st. Sensitivity to albedo timing and range can lead to significant changes in energy balance calculations, and accounting for these changes may improve long-term climate models (Sellers, et al., 1995) (Davidson, et al., 2004). As a result of this sensitivity, investigations into the long-term impacts of evolving cropping pattern effects on atmospheric changes should also consider effects of the overall albedo shift caused by the changes in the growth profile and reflectivity of each crop.

4.3.3 Uncertainty analysis

As mentioned in the dataset section, CDL data, which are constructed using satellite observations from satellites such as RESOURCESAT-1 and Landsat, contain uncertainties. Provided with the CDL dataset is the uncertainty of each classified crop type in the form of accuracy assessments. Based on these NASS accuracy assessments the overall user accuracy of the CDL was calculated to be ~80% on average for all land cover types each year; however, the major croplands utilized in this study generally fared better, with user accuracies averaging near 90% annually (as shown in popular US crops listed in Table 4). To minimize the impact of this uncertainty in crop type classification, only MODIS pixels that contained completely homogenous CDL croplands were

selected for analysis. Areas that appear noisy due to the difficulty of classification, such as field edges or variable wetlands, were also excluded. In addition, due to the extensive replication of both points and years, effects from erroneous classifications would be reduced in the final albedo calculations due to the sheer number of correctly classified croplands incorporated into the analysis.

Table 4: National CDL User Accuracy of Popular Crops from 2015-2018, note that in 2015 user accuracy was calculated using a buffered method, while 2016-2018 was calculated through an unbuffered method (source: (USDA-NASS, 2020)).

	Total Count				User Accuracy			
	2015	**2016**	**2017**	**2018**	**2015***	**2016**	**2017**	**2018**
Soybeans	4,579,903	4,618,482	5,022,357	4,945,568	93%	91%	92%	92%
Maize	4,651,997	4,839,604	4,639,961	4,713,799	95%	92%	93%	94%
Winter Wheat	2,165,988	2,030,583	1,730,739	1,504,814	91%	89%	88%	87%
Other Hay/Non Alfalfa	1,270,865	1,596,983	1,689,644	1,518,819	78%	76%	80%	81%
Alfalfa	1,241,323	1,380,339	1,419,039	1,288,739	87%	84%	84%	85%
Cotton	783,641	632,168	705,317	1,105,499	85%	81%	84%	86%
Spring Wheat	539,928	451,067	472,286	472,955	87%	84%	82%	84%

For the MODIS BRDF albedo product used in the study, evaluation efforts were undertaken to determine the validity of both the broadband and individual albedo channels under multiple conditions and testing regimes ((Disney, et al., 2004), (Jin, et al., 2003), (Knobelspiesse, et al., 2008)), which found general agreement of other measurement techniques with the MODIS albedo. Still, it is possible that regions with the presence of thin clouds and aerosol layers may be misidentified as clear regions, as MODIS is insensitive to very optically thin cirrus clouds with optical depth less than 0.3 ((Huang, et al., 2011) (Minnis, et al., 2008)). It is worth noting that the MODIS albedo product from the MODIS MAIAC data (Lyapustin, et al., 2012) is also available.

However, the MODIS MAIAC albedo product is only available every 8 days at a spatial resolution of 1 km. Yet, typical field sizes in the United States do not exceed 0.5 miles per side, effectively limiting the resolution to 800m or better for an albedo dataset to encompass MODIS pixels within single, homogeneous fields. Due to this spatial resolution limitation, as well as the reduced temporal resolution of 8 days between albedo calculations, the albedo product from the MAIAC data is not used.

Another area of concern is the level-3 nature of the MODIS BDRF product, which features a 16-day weighted method to estimate the albedo on the middle-most day of the sequence. This provides a continuous albedo curve for the entire reference period but can serve to smooth out both peaks and valleys of the albedo function, reducing accuracy at those transition points during the growing season. However, the advantage of using level-3 gridded data is consistent size and location of the derived pixels, which allows the field locations to be constant throughout the study period, eliminating possible uncertainty arising from a daily changing pixel size and location as we would see with similar level-1, non-gridded, datasets.

Finally, while this study focused on crops in the United States, varieties and management practices vary greatly on a global scale. For example, planting date and irrigation practices of maize in regions that often double crop (such as those seen in the southern Asia region) can vary greatly from similar climatological areas found in the United States. Additionally, each individual crop can have several varieties, for example, as of February 2020 the USDA's Plant Variety Protection office lists 2134 varieties of field corn, 2752 varieties of soybean, and 1294 varieties of common wheat (USDA-AMS,

2020). Due to the differences in both climate and management practices these varieties can vary in both growth speed, stage timing, and overall leaf area index impacting the timing and scope of albedo changes throughout the season. As such, albedo changes derived from varieties found in the US may not be representative of albedo changes cause by varieties found outside the study area, even when factoring in planting dates, management, and weather conditions. As such studies into the specific region's crop growth patterns and timings should be completed to verify albedo patterns are similar to US results before incorporating the results from this database into studies outside this study's scope.

4.3.4 Evaluating the feasibility of using changes in NDVI as a proxy for changes in albedo over cropland

In some of the previous studies, the changes in albedo over cropland during the growing seasons are approximated using NDVI values (e.g. (Hsu, et al., 2013)). The R^2 values between NDVI and black sky albedo values at seven wavelengths (470nm, 555nm, 645nm, 860nm, 1240nm, 1640nm, and 2130nm) are studied (Figure 22) using 4 years data as mentioned in Section 4.2. On an annual basis, a maximum R^2 value of 0.75 is found at 650 nm and the minimum R^2 value of 0.2 is found at 1240nm. Also as shown in Figure 22, this correlation on a daily scale increases over cropland from near zero correlation early in the growing season, reaching maximum correlation in the late peak growth season, and decreasing shortly thereafter back to near zero as crops are removed from the surface during harvest. This exercise suggests that NDVI may not be an ideal

proxy for capturing albedo changes over cropland during the growing seasons and, as a result, further investigation into the impacts of surface crop type and land-use mixture on remote sensing applications may be warranted.

Figure 22: Coefficient of determination for each wavelength's albedo using NDVI over United States cropland areas for 2015-2018 by day of year (colored dots) and annual total (black dots).

4.4 Conclusion

Using four-year (2015-2018) of CDL and MODIS BRDF data, the changes in black- and white-sky albedo at seven visible to Shortwave IR wavelengths (470nm, 555nm, 645nm, 860nm, 1240nm, 1640nm, 2130nm) as well as the broadband SW (300-5000nm) during the planting, growing and harvesting periods were studied as functions of crop type and plant hardiness zones (HZ) over the continental United States. A total

of 55 types of crop were included in the study, with an emphasis on four commonly planted crops: maize, soybean, spring wheat and cotton. This study found:

1. Significant change in black-sky surface albedo at all seven channels are found between planting, peak growing, and harvest period for maize, soybean, and spring wheat. A total of a 70% decrease in 645 nm channel albedo is found for maize and soybean from early planting to the peak of the growing season, similar but less significant changes on the order of 60% and 30% are also found for spring wheat and cotton respectively. Additionally, similar but less significant changes are also found for the blue (470 nm) and green (555nm) channels.

2. For 1640 and 2140nm spectral channels, although the overall black-sky albedo is much higher compared with the visible channels, lower albedo values are also observed at the peak growing period, with a ~0.1 lower albedo found for the planting season for all four types of crops. The opposite pattern, however, is found for the 860 nm channels, with the peak surface albedo found during the peak growing season with a ~0.2 albedo rise, compared to the planting season for maize, soybean, and spring wheat. A similar pattern, albeit weak in magnitude is found for cotton fields. Also note that the white-sky albedo behavior is nearly indistinguishable to the black sky albedo and therefore are not mentioned hereafter.

3. An interesting pattern is found for the 1240 nm spectrum channel, with lower black-sky albedo values found for the planting seasons and higher black-sky albedo values found in both peak growing and harvesting seasons for maize, soybean, and spring wheat. The reason for this observed pattern is not known, but may be due to the closeness

of the soil/plant reflectivity coupled with the sensitivity to shallow surface water conditions at this channel.

4. In the broadband shortwave, both crop type and growth cycle were found to affect the overall black-sky albedo, varying by 40% relative (0.04 absolute) to the black-sky albedo at planting for maize, soybeans, and spring wheat, with cotton black-sky albedo changing only slightly, 25% relative (0.02 absolute). The significant variation in broadband albedo should be taken into consideration in climate applications over croplands in future studies.

5. We also found that NDVI may not be a good proxy for simulating cropland albedo changes during the growing cycle, as the R^2 values range from 0.2 to 0.75 between NDVI and black-sky surface albedo, for the seven spectral channels used in this study; including the key 555nm channel which has an R^2 value of approximately 0.4.

Lastly, the goal of the study is to construct a cropland surface albedo database relating plant type, hardiness zone, and day of year to expected surface albedo using averages of nationwide, single-crop MODIS BRDF values over a four-year period (2015-2018) for a total of 55 crop types. This database provides the means to calculate appropriate albedo parameters for albedo sensitive applications, such as long-term weather modeling, for the included cropland variations within the United States. The database includes the mean, standard deviation, and pixel counts for matching crop, location, and time conditions for both white- and black-sky albedo allowing for a broad use of the data across multiple types of studies. For example, linking the resulting albedo parameters to economic modelling of crop acreage response functions could allow estimation of currently

unquantified surface albedo and regional climate effects of US agricultural and trade policies that affect farmers' choice of crops via the price mechanism. The database is attached as a supplement for this study for potential remote sensing, weather, and climate applications that require accurate quantification of lower boundary conditions.

CHAPTER 5

SUMMARY AND CONCLUSIONS

In this thesis two different studies were conducted, centered on crop land and agricultural related efforts, for studying the impacts of economics factors on agricultural activities as well as the changes in narrowband and broad band albedo over the crop land during the plant growing period that could impact a variety of applications in remote sensing and atmospheric modeling.

Through the two-way coupling of an economics model with the ALMANAC crop simulation model we found the prairie pothole region of North Dakota could be simulated under multiple economics scenarios. This is a one of its kind proof-of concept study that aims for accounting for economic factors in crop simulations. This study suggests that it is feasible to incorporate economics factors in crop simulations to account for economical-induced land use change due to crop rotation and enabled the exploration of policies, such as grassland incentives, and market influences, such as crop price surges, at a fine scale over a large area.

Additionally, crop land albedo varies significantly during the plant growing period that can impact applications over crop land regions that using crop land surface albedo as one of the low boundary conditions, such as remote sensing and aerosol modeling applications. Through the evaluation of the MODIS BRDF product collocated with the CDL, we constructed a database of both channel and broadband albedo values found separated into individual crops, HZs, and day of year. This

database will allow future studies and models to incorporate specific crop albedos throughout the growing season with a minimal impact to computational resources.

References

Boryan, Claire, Zhengwei Yang, Rick Mueller, and Mike Craig. 2011. "Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program." *Geocarto International* 341-358.

Boryan, Claire, Zhengwei Yang, Rick Mueller, and Mike Craig. 2011. "Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program." *Geocarto International* 341-358. doi:10.1080/10106049.2011.562309.

Chappelle, Emmett W., Moon S. Kim, and James E. McMurtrey III. 1992. "Ratio Analysis of Reflectance Spectra (RARS): An Algorithm for the Remote Estimation of the Concentrations of Chlorophyll A, Chlorophyll B, and Carotenoids in Soybean Leaves." *Remote Sensing of Environment* 239-247. doi:10.1016/0034-4257(92)90089-3.

Daly, Christopher, George Taylor, and Wayne Gibson. 1997. "The Prism Approach to Mapping Precipitation and Temperature." *10th AMS Conference on Applied Climatology.* Reno, NV.

Daly, Christopher, Mark P. Widrlechner, Michael D. Halbleib, Joseph I. Smith, and Wayne P. Gibson. 2012. "Development of a New USDA Plant Hardiness Zone Map for the United States." *Journal of Applied Meteorology and Climatology* 242-264. doi:10.1175/2010JAMC2536.1.

Davidson, Andrew, and Shusen Wang. 2004. "The effects of sampling resolution on the surface albedos of dominant land cover types in the North American boreal region." *Remote Sensing of Environment* 211-224. doi:10.1016/j.rse.2004.07.005.

Disney, Mathias I., Philip E. Lewis, Graham Thackrah, Tristan L. Quaife, and Michael J. Barnsley. 2004. "Comparison of MODIS broadband albedo over an agricultural site with ground measurements and values derived from earth observation data at a range of spatial scales." *International Journal of Remote Sensing* 5297-5317. doi:10.1080/01431160410001720180.

Hsu, N. C., Myeong-Jae Jeong, Corey Bettenhausen, A. M. Sayer, Richard A. Hansell, Colin Seftor, Jin Huang, and Si-Chee Tsay. 2013. "Enhanced Deep Blue aerosol retrieval algorithm: The second generation." *Journal of Geophysical Research: Atmospheres* 9296–9315. doi:10.1002/jgrd.50712.

Huang, Jingfeng, N. Christina Hsu, Si-Chee Tsay, Myeong-Jae Jeong, Brent N. Holben, Timothy A. Berkoff, and Ellsworth J. Welton. 2011. "Susceptibility of aerosol optical thickness retrievals to thin cirrus contamination during the BASE-ASIA campaign." *Journal of Geophysical Research.* doi:10.1029/2010JD014910.

Jandaghian, Zahra, and Umberto Berardi. 2020. "Analysis of the cooling effects of higher albedo surfaces during heat waves coupling the Weather Research and Forecasting model with building energy models." *Energy & Buildings.* doi:10.1016/j.enbuild.2019.109627.

Jin, Yufang, Crystal B. Schaaf, Feng Gao, Xiaowen Li, Alan H. Strahler, Wolfgang Lucht, and Shunlin Liang. 2003. "Consistency of MODIS surface bidirectional reflectance distribution function and albedo retrievals: 1. Algorithm

performance." *Journal of Geophysical Research Atmospheres.* doi:10.1029/2002JD002803.
Keating, B.A., P.S. Carberry, G.L. Hammer, M.E. Probert, M.J. Robertson, D. Holzworth, N.I. Huth, et al. 2003. "An overview of APSIM, a model designed for farming systems simulation." *European Journal of Agronomy* 18: 267-288.
Kharel, Gehendra, Haochi Zheng, and Andrei Kirilenko. 2016. "Can land-use change mitigate long-term flood risks in the Prairie Pothole Region? The case of Devils Lake, North Dakota, USA." *Regional Environmental Change* 2443-2456.
Kiniry, J. R., J. R. Williams, P. W. Gassman, and P. Debaeke. 1992. "A General, Process-Oriented Model for Two Competing Plant Species." *Transactions of the ASAE* 801-810.
Knobelspiesse, Kirk D., Brian Cairns, Beat Schmid, Miguel O. Roman, and Crystal B. Schaaf. 2008. "Surface BRDF estimation from an aircraft compared to MODIS and ground estimates at the Southern Great Plains site." *Journal of Geophysical Research Atmospheres.* doi:10.1029/2008JD010062.
Lucht, Wolfgang, Crystal Barker Schaaf, and Alan H. Strahler. 2000. "An Algorithm for the Retrieval of Albedo from Space Using Semiempirical BRDF Models." *IEEE Transactions on Geoscience and Remote Sensing* 977-998. doi:10.1109/36.841980.
Lyapustin, Alexei I., Yujie Wang, Istvan Laszlo, Thomas Hilker, Forrest G.Hall, Piers J. Sellers, Compton J. Tucker, and Sergey V. Korkin. 2012. "Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction." *Remote Sensing of Environment* 385-393. doi:10.1016/j.rse.2012.09.002.
Mearns, L. O., T. Mavromatis, and E. Tsvetsinska. 1999. "Comparative responses of EPIC and CERES crop models to high and low spatial resolution climate change scenarios." *Journal of Geophysical Research* 6623-6646.
Mesinger, Fedor, Geoff DiMego, Eugenia Kalnay, Kenneth Mitchell, Perry C. Shafran, Wesley Ebisuzaki, Dušan Jović, et al. 2006. "North American Regional Reanalysis." *Bulletin of the American Meteorological Society* 343-360.
Minnis, Patrick, Qing Z. Trepte, Szedung Sun-Mack, Yan Chen, David R. Doelling, David F. Young, Douglas A. Spangenberg, et al. 2008. "Cloud Detection in Nonpolar Regions for CERES Using TRMM VIRS and Terra and Aqua MODIS Data." *IEEE Transactions on Geoscience and Remote Sensing* 3857-3884. doi:10.1109/TGRS.2008.2001351.
Richards, John A. 2013. *Remote Sensing Digital Image Analysis.* Berlin: Springer. doi:10.1007/978-3-642-30062-2.
Ritchie, Joe, and S. Otter. 1985. *Description and performance of CERES-Wheat: A user-oriented wheat yield model.* United States Department of Agriculture, Agricultural Research Service.
Rouse Jr, J.W., R.H. Haas, J.A. Schell, and D.W. Deering. 1973. "Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation."

Schaaf, Crystal B., Feng Gao, Alan H. Strahler, Wolfgang Lucht, Xiaowen Li, Trevor Tsang, Nicholas C. Strugnell, et al. 2002. "First operational BRDF, albedo nadir reflectance products from MODIS." *Remote Sensing of Environment* 135-148. doi:10.1016/S0034-4257(02)00091-3.

Seidel, Felix, Alexander Kokhanovsky, and Michael Schaepman. 2012. "Fast retrieval of aerosol optical depth and its sensitivity to surface albedo using remote sensing data." *Atmospheric Research* 22-32. doi:10.1016/j.atmosres.2011.03.006.

Sellers, P. J., B. W. Meeson, F. G. Hall, G. Asrar, R. E. Murphy, R. A. Schiffer, F. P. Bretherton, et al. 1995. "Remote Sensing of the Land Surface for Studies of Global Change: Models Algorithms Experiments." *Remote Sensing Environment* 3-26. doi:10.1016/0034-4257(94)00061-Q.

Soil Conservation Service, USDA. 1961. *Land-Capability Classification.* Washington, DC: U.S. Government Printing Office.

Soil Survey Staff. 2016. *Natural Resources Conservation Service, United States Department of Agriculture, Soil Survey Geographic (SSURGO) Database.* Available online at https://sdmdataaccess.sc.egov.usda.gov.

USDA - NASS. 2020. *Acreage.* United States Department of Agriculture.

USDA - NASS. 2020. *Farms and Land in Farms, February 2020.* United States Department of Agriculture.

USDA - NASS. 2010. "Field Crops Usual Planting and Harvesting Dates."

USDA-AMS. 2020. "Scanned and Redacted Issued Certificates." *Agricultural Marketing Service: Plant Variety Protection Office.* 02 14. https://apps.ams.usda.gov/CMS/.

USDA-ARS. 2012. "USDA Plant Hardiness Zone Map." *Agricultural Research Service.* Accessed 2019. https://planthardiness.ars.usda.gov/.

USDA-NASS. 2016. *Cropland Data Layer.* Available at https://nassgeodata.gmu.edu/CropScape/.

—. 2018. *NASS QuickStats.* https://quickstats.nass.usda.gov/.

Vermote, Eric. 2015. *MYD09 MODIS/Aqua L2 Surface Reflectance, 5-Min Swath 250m, 500m, and 1km.* doi:10.5067/MODIS/MYD09.006.

White, Jeffrey W., Gerrit Hoogenboom, Bruce A. Kimball, and Gerard W. Wall. 2011. "Methodologies for simulating impacts of climate change on crop production." *Field Crops Research* 124.

Williams, J. R., K. G. Renard, and P. T. Dyke. 1983. "Viewpoint: EPIC: A new method for assessing erosion's effect on soil productivity." *Journal of Soil and Water Conservation* 381-383.

Xie, Yun, James R. Kiniry, Vernon Nedbalek, and Wesley D. Rosenthal. 2001. "Maize and Sorghum Simulations with CERES-Maize, SORKAM, and ALMANAC under Water-Limiting Conditions." *Agronomy Journal* 93: 1148–1155.

Zelinka, Mark D., Timothy A. Myers, Daniel T. McCoy, Stephen Po-Chedley, Peter M. Caldwell, Paulo Ceppi, Stephen A. Klein, and Karl E. Taylor. 2020. "Causes of Higher Climate Sensitivity in CMIP6 Models." *Geophysical Research Letters.* doi:10.1029/2019GL085782.